REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE 11MAR03	3. REPORT TYPE AND DATES COVERED FINAL, 01MAR99 – 31MAY02 June

4. TITLE AND SUBTITLE
Integrated Microelectronics and Photonics Active Cooling Technology (IMPACT)

5. FUNDING NUMBERS
DAAD19-99-1-0158

6. AUTHOR(S)
J. Bowers, A. Shakouri, A. Majumdar, V. Narayanamurti, E. Croke

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
John Bowers, University of California at Santa Barbara, ECE Department
Santa Barbara, CA 93106

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
P-40068-EL
.1

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The primary goal of the IMPACT effort is to demonstrate the advantages of heterostructure integrated thermionic (HIT) coolers and their integration with microelectronics and photonics. The majority of our research involves the development of this new technology through nanostructured materials design and growth; device design and fabrication; simulation and modeling; novel measurements of thermoelectric and thermionic behavior; and systems integration and packaging.

14. SUBJECT TERMS
cooling technology, thermionic coolers, thermoelectric coolers, semiconductor heterostructures, integrated microelectronics and photonics, nanostructures, packaging

15. NUMBER OF PAGES
135

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev.2-89)
Prescribed by ANSI Std. 239-18
298-102

Integrated Microelectronics and Photonics Active Cooling Technology (IMPACT)

Final Report

Performers: University of California Santa Barbara
University of California Santa Cruz
University of California Berkeley
Harvard University
HRL Laboratories

Professor John Bowers, UCSB (Director) (805) 893-8447
Professor Ali Shakouri, UCSC (Technical Director) (831) 459-3821

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

Agent: Army Research Office
Dwight Woolard (919) 549-4297

Table of Contents

Abstract

The primary goal of the IMPACT effort is to demonstrate the advantages of heterostructure integrated thermionic (HIT) coolers and their integration with microelectronics and photonics. The majority of our research involves the development of this new technology through nanostructured materials design and growth; device design and fabrication; simulation and modeling; novel measurements of thermoelectric and thermionic behavior; and systems integration and packaging.

During our effort, SiGe/Si superlattice microcoolers have been integrated with thin film heaters. Use of a heat load on top of device allowed direct measurement of cooling power density. It was observed that smaller size cooler devices (40-50 micron in diameter) have cooling power density much larger than larger devices (100 micron in diameter). A detailed theory has been developed which predict accurately the maximum cooling temperature and cooling power density for various device sizes. We improved the characterization of cross sectional temperature profile in thin film thermionic coolers with <100nm nm resolution using scanning thermal microscopy. By modulating the device current at a known frequency, ω, the Peltier cooling/heating contribution (proportional to ω) and Joule heating contribution (proportional to 2ω) were separated. Using thermoreflectance imaging technique, frequency response of micro coolers were measured for several devices with different superlattice periods. It was found that the frequency response (>40-50kHz) is limited by the thermal mass of the metalization on top of the device and it is independent of the superlattice thickness or the device diameter. We continued 3-ω measurements of thermal conductivity (77-400K) of various superlattice structures in order to minimize the thermal conductivity. Electron transmission in various superlattices has also been experimentally studied using ballistic electron emission microscopy.

Scientific Personnel

John E. Bowers	**(UCSB)**	**PI**
Chris Labounty	(UCSB)	PhD (earned under Heretic)
Xiaofeng Fan	(UCSB)	PhD (earned under Heretic)
Gehong Zeng	(UCSB)	Visiting Researcher
Ali Shakouri	**(UCSC)**	**Co-PI**
Daryoosh Vashaee	(UCSC)	PhD Student
Yan Zhang	(UCSC)	PhD Student
James Christofferson	(UCSC)	MS Student
Arun Majumdar	**(UCB)**	**Co-PI**
Andrew Miner	(UCB)	PhD Student
Scott Huxtable	(UCB)	PhD Student
Venky Narayanamurti	**(Harvard)**	**Co-PI**
R.G. Mani	(Harvard)	Research Associates
I. Altfeder	(Harvard)	Research Associates
J. Yoon	(Harvard)	PhD Student
Ed Croke	**(HRL)**	**Co-PI**
Howard Dunlap	(HRL)	Researcher
Kevin Holabird	(HRL)	Researcher

Scientific Progress and Accomplishments

Project Goals

The primary goal of the IMPACT effort is to demonstrate the advantages of heterostructure integrated thermionic (HIT) coolers and their integration with microelectronics and photonics. The majority of our research involves the development of this new technology through nanostructured materials design and growth; device design and fabrication; simulation and modeling; novel measurements of thermoelectric and thermionic behavior; and systems integration and packaging.

Approach

Materials design is focused on increasing the cooling power and efficiency with thermionics and phonon bandgap engineering in superlattices. Electrical and thermal transport measurements are used to verify model predictions and aid in further improvements in device and materials design. Simulations are used to determine device limitations and non-ideal effects. Ultimately, thermionic devices are to be integrated and packaged for systems demonstration.

Accomplished Milestones

- Design and fabrication of both n- and p-SiGe/Si superlattice coolers on a SOI (silicon on insulator) substrate.
- Detailed 2D and 3D electro thermal modeling of superlattice coolers. Accurate prediction of maximum cooling (4.5K at room temperature) and cooling power density (680W/cm2) for various device sizes and superlattice material.
- Integration of micro thin film heaters with cooler devices to characterize the cooling temperature and cooling power density of different superlattices thicknesses.
- Measured frequency response of SiGe superlattice micro coolers using thermoreflectance imaging technique (>40-50kHz for superlattice thickness 1-6um, device size 40-100um in diameter).
- 3-ω measurements of thermal conductivity (77-400K) of various superlattice structures.
- Cross sectional thermal microscopy of thin film thermionic coolers with <100nm nm resolution (separation of Peltier and Joule heating contributions).
- Substrate transfer of thin films to copper substrates for improved heat sink properties.
- Fabrication and characterization of SiGe superlattice microcooler samples on high thermal conductivity single-isotope Si substrate.

- Monolithic integration of InGaAsP/InP based thin film coolers with in-plane laser structures emitting at 1.55 μm.
- Ballistic electron emission microscopy of various InAlP/InGaP superlattices. The effects of momentum randomization at a metal/semiconductor interface have also been measured.
- Further development of n- and p-type superlattice cooler designs for future integration of p- and n-type elements.
- 3-ω measurements of thermal conductivity (77-400K) of highly strained Si0.1Ge0.9/Si0.9Ge0.1 superlattice structures.
- Develop microcooler arrays for multi-place localized temperature control in a single chip.

Awards

Ali Shakouri:
NSF CAREER Award April 2000
Packard Fellowship 1999

V. Narayanamurti:
Centennial Plenary Speaker, March Meeting of the American Physical Society, "Frontiers in Condensed Matter and Materials Physics."

Chair, National Research Council, Board of Physics and Astronomy panel on "Condensed Matter and Materials Physics; Basic Research for Tomorrow's Technology."

Year 2000 Individual Academic Achievement Award at the Asian American Unity Dinner, Boston, May, 2000.

In March 2000, the Venky Narayanamurti Entrepreneurship Award was endowed at the University of California, Santa Barbara.

Inventions

1. Ali Shakouri, James Christofferson, and Peyman Milanfar, "Sub micron thermal imaging", UC Case No. 2001-161-1, October. 2000, (Patent pending)
2. Peyman Milanfar and Ali Shakouri, "Superresolution with AC imaging," UC Case No. 2002-047, July 2001 (patent pending)
3. Ken Pedrotti and Ali Shakouri, "Differential Boxcar Imaging Chip," Record of Invention submitted to University of California, October 200

Publications

(a) Published in peer-reviewed journals (attached)

Chen, G.; Shakouri, A. "Heat transfer in nanostructures for solid-state energy conversion," Transactions of the ASME. Journal of Heat Transfer, vol.124, (no.2), ASME, April 2002. p.242-52.

Huxtable, S.T.; Abramson, A.R.; Chang-Lin Tien; Majumdar, A.; LaBounty, C.; Xiaofeng Fan; Gehong Zeng; Bowers, J.E.; Shakouri, A.; Croke, E.T. "Thermal conductivity of Si/SiGe and SiGe/SiGe superlattices," Applied Physics Letters, vol.80, (no.10), AIP, 11 March 2002. p.1737-9.

R.E. MartinezII, I. Appelbaum, C.V. Reddy, R. Sheth, K.J. Russel and V. Narayanamurti, "Electron Transport through strongly coupled AlInP/GnInP Superlatives", Applied Physics Letters, vol. 81, p. 3576 (2002)

"Quantum-Confinement-Assisted Electron Transport Across Metal Semiconductor Interface," by Igor Altfeder, J. A. Golovchenko and V. Narayanamurti, Phys. Rev. Lett., Vol. 87, No. 5, 056801 (2001).

"BEEM imaging and spectroscopy of buried structures in semiconductors," by V. Narayanmurti and M. Kozhevnikov, Physics Reports 34916, 447-514(2001) Invited Review

C. LaBounty, A. Shakouri, and J. E. Bowers, "Design and Characterization of Thin Film Micro-Coolers," Journal of Applied Physics, Vol. 89, No. 7, 1 April 2001.

X. Fan, G. Zeng, C. LaBounty, E.Croke, A. Shakouri, C. Ahn, J.E. Bowers, "SiGeC/Si superlattice micro cooler," Applied Physics Lett., Vol. 78, No.11, 12 March 2001.

X. Fan, G. Zeng, C. LaBounty, E.Croke, D. Vashaee, A. Shakouri, C. Ahn, J.E. Bowers, "High cooling power density SiGe/Si micro coolers," Elec. Lett., Vol. 37, No. 2, 18th January 2001.

C. LaBounty, A. Shakouri, P. Abraham, and J. E. Bowers, "Monolithic integration of thin film coolers with optoelectronic devices," *Optical Engineering,*Vol. 39, No.11, November 2000.

X. Fan, G. Zeng, E. Croke, G. Robinson, C. LaBounty, A. Shakouri, J.E. Bowers, "SiGe/Si Superlattice Coolers," Phys. Low-Dim. Struct., 5/6 (2000) pp. 1-10.

S. T. Huxtable, C. LaBounty, A. Shakouri, P. Abraham, Y. J. Chiu, X. Fan, J. E. Bowers, and A. Majumdar, "Thermal Conductivity of Indium Phosphide Based Superlattices," *Microscale Thermophysical Engineering,* Vol.4, No. 3 (2000).

(b) Published in conference proceedings

Zhang, Y.; Zeng, G.; Singh, R.; Christofferson, J.; Croke, E.; Bowers, J.E.; Shakouri, A. "Measurement of Seebeck coefficient perpendicular to SiGe superlattice layers," 21st International Conference on Thermoelectrics, Long Beach CA, 26-29 August 2002.

A. Shakouri, D. Vashaee, J. Christofferson, X. Fan and C. Labounty, "Thin film solid state refrigerators for spot cooling of integrated circuits," Thermal Challenges in Next Generation Electronic Systems (THERMES 2002), Edited by Y.K. Joshi and S. V. Garimella, pp.75-76, January 2002.

A. Fitting, J. Christofferson, A. Shakouri, X. Fan, G. Zeng, C. LaBounty, J.E. Bowers, E. Croke "Transient response of thin film SiGe micro refrigerators" 2001 ASME International Mechanical Engineering Congress and Exposition November 11-16, 2001, New York, NY

A. Majumdar, S. Huxtable, A. Abramson, C. Tien, X. Fan, G. Zeng, C. LaBounty, J.E. Bowers, E. Croke, A. Shakouri "Nanoscale thermal transport in superlattices" 2001 ASME International Mechanical Engineering Congress and Exposition November 11-16, 2001, New York, NY

A. Majumdar, K. Tan, A. Miner, X. Fan, G. Zeng, C. LaBounty, J.E. Bowers, E. Croke, A. Shakouri, "Nanoscale thermal imaging of thermionic devices using scanning thermal microscopy" 2001 ASME International Mechanical Engineering Congress and Exposition November 11-16, 2001, New York, NY

J. Christofferson, D. Vashaee, A. Shakouri, P. Melese "Real time submicron imaging using thermoreflectance" 2001 ASME International Mechanical Engineering Congress and Exposition November 11-16, 2001, New York, NY

C. LaBounty, X. Fan, G. Zeng, Y. Okuno, A. Shakouri, and J.E. Bowers, "Integrated Cooling of Semiconductor Lasers by Wafer Fusion," 20th International Conference on Thermoelectrics, Beijing, China, June 2001.

Xiaofeng Fan, Gehong Zeng, Edward Croke, Chris LaBounty, Daryoosh Vashaee, James Christofferson, Ali Shakouri, and John E. Bowers, "Integrated cooling for Si-based microelectronics", Twentieth International Conference on Thermoelectrics, ICT 2001.

J. Christofferson, D. Vashaee, A. Shakouri, P. Melese, X. Fan, G. Zeng, C. LaBounty, and J.E. Bowers, "Thermoreflectance Imaging of Superlattice Micro Refrigerators," SEMITHERM XVII symposium proceedings, San Jose, Ca, March 2001.

James Christofferson, Daryoosh Vashaee, Ali Shakouri, Phil Melese, "Opto-thermal characterization of heterostructure integrated thermionic coolers", SPIE-Photonic West Metrology-Based Control for Micro Manufacturing, San Jose, Ca, January 2001

Daryoosh Vashaee, Christopher LaBounty, Xiaofeng Fan, Gehong Zeng, Patrick Abraham, John E. Bowers, and Ali Shakouri, "Modeling of Thin film P-InGaAs/InGaAsP coolers for Integrated Optical Devices", Photonics West, 20-26 January 2001, San Jose, CA

C. LaBounty, D. Oberle, J. Piprek, P. Abraham, A. Shakouri, J.E. Bowers, "Monolithic Integration of Solid State Thermionic Coolers with Semiconductor Lasers," *LEOS 2000 Conference Proceedings (IEEE)*, Rio Grande, Puerto Rico, November 2000.

C. LaBounty, G. Robinson, P. Abraham, A. Shakouri, J.E. Bowers, "Transferred-Substrate InGaAsP-based thermionic emission coolers," *IMECE 2000 Conference Proceedings (ASME)*, Orlando, FL, November 2000.

J. Christofferson, D. Vashaee, A. Shakouri, X. Fan, G. Zeng, C. LaBounty, J.E. Bowers, E.T. Croke III, "Thermal Characterization of Thin Film Superlattice Micro Refrigerators," *International Symposum on Compound Semiconductors*, Monterey, CA, September 2000.

C. LaBounty, D. Vashaee, X. Fan, G. Zeng, P. Abraham, A. Shakouri, J.E. Bowers ,"P-type InGaAs/InGaAsP Superlattice Coolers," *19th International Conference on Thermoelectrics*, Cardiff, Wales, August 2000.

C. LaBounty, A. Shakouri, P. Abraham, and J. E. Bowers, "Two Stage Monolithic Thin Film Coolers," ITherm 2000 Conference Proceedings, Las Vegas, NV, May 2000.

X. Fan, G. Zeng, E. Croke, G. Robinson, C. LaBounty, A. Shakouri, and J.E. Bowers, "n- and p-type SiGe/Si Superlattice Coolers," ITherm 2000 Conference Proceedings, Las Vegas, NV, May 2000.

C. LaBounty, A. Shakouri, G. Robinson, L. Esparza, P. Abraham, and J. E. Bowers, "Experimental Investigation of Thin Film Coolers," *Material Research Society Spring Meeting*, San Francisco, CA, April 2000.

X. Fan, G. Zeng, E. Croke, G. Robinson, C. LaBounty, A. Shakouri, and J.E. Bowers, "P SiGe/Si Superlattice Cooler," *Material Research Society Spring Meeting*, San Francisco, CA, April 2000.

Gehong Zeng; Shakouri, A.; Bounty, C.L.; Robinson, G.; Croke, E.; Abraham, P.; Xiafeng Fan; Reese, H.; Bowers, J.E. "SiGe micro-cooler," Electronics Letters, vol.35 (24), 25 Nov. 1999. p.2146-7

(c) Presented (Abstracts):

Scott T. Huxtable, Alexis R. Abramson, Arun Majumdar, Ali Shakouri, and Edward T. Croke "The effect of defects and acoustic impedance mismatch on heat conduction in SiGe based superlattices," International Mechanical Engineering Congress and Exhibition (IMECE 2002), New Orleans, LA, Nov. 2002 (to be published)

C. LaBounty, D. Vashaee, J. Christofferson, X. Fan, G. Zeng, A. Shakouri, J.E. Bowers, "Recent advances in InP-based thermionic coolers," American Physical Society March Meeting 2001 in Seattle, WA, March 12-16, 2001.

D. Vashaee, A. Shakouri, "Hot electron effects in heterostructure integrated thermionic coolers," American Physical Society March Meeting 2001 in Seattle, WA, March 12-16, 2001.

X. Fan, G. Zeng, C. LaBounty, E.Croke, A. Shakouri, J.E. Bowers, "The effect of superlattice pararmeters on the performance fo SiGe/Si thin film coolers," American Physical Society March Meeting 2001 in Seattle, WA, March 12-16, 2001.

"Ballistic Electron Emission Spectroscopy on GaAs/AlGaAs Superlattices," by Joonah Yoon, V. Narayanamurti and Yi-Jen Chiu. APS March meeting in Seattle, 2001

"Ballistic Electron Emission Spectroscopy on GaAs/AlGaAs Superlattices," by Joonah Yoon, V. Narayanamurti and Yi-Jen Chiu. Paper presented at the MRS meeting, Boston, MA, December 2000 and to be presented at the APS March meeting in Seattle, 2001

C. LaBounty, G. Almuneau, A. Shakouri, J.E. Bowers, "Sb-based III-V Cooler," *19th International Conference on Thermoelectrics*, Cardiff, Wales, August 2000.

X. Fan, G. Zeng, C. Luo, E. Croke, D. Clarke, A. Shakouri, and J.E. Bowers, "Thermal Conductivity of SiGe/Si and SiGeC/Si Superlattices and their Application to Thermoelectric/Thermionic Cooling", Fourteenth Symposium on Thermophysical Properties, Boulder, CO, June 2000.

"Ballistic Electron Emission Microscopy (BEEM) of Au/Al/GaAs Schottky barriers under Ultra High Vacuum Conditions," R.G. Mani and V. Narayanamurti, Bulletin American Physical Society, 45, p. 848, (2000)

Heat Transfer in Nanostructures for Solid-State Energy Conversion

G. Chen
Assoc. Prof. Mem. ASME
Mechanical Engineering Department,
Massachusetts Institute of Technology,
Cambridge, MA 02139

A. Shakouri
Jack Baskin School of Engineering,
University of California,
Santa Cruz, CA 95064-1077

Solid-state energy conversion technologies such as thermoelectric and thermionic refrigeration and power generation require materials with low thermal conductivity but good electrical conductivity and Seebeck coefficient, which are difficult to realize in bulk semiconductors. Nanostructures such as superlattices, quantum wires, and quantum dots provide alternative approaches to improve the solid-state energy conversion efficiency through size and interface effects on the electron and phonon transport. In this review, we discuss recent research and progress using nanostructures for solid-state energy conversion. The emphasis is placed on fundamental issues that distinguish energy transport and conversion between nanoscale and macroscale, as well as heat transfer issues related to device development and property characterization. [DOI: 10.1115/1.1448331]

Keywords: Conduction, Energy Conversion, Heat Transfer, Microscale Microstructures, Nanoscale, Thermoelectric, Thin Films

1 Introduction

Accompanying the motion of charges in conductors or semiconductors, there is also an associated energy transport. Consider a current flowing through a pair of n-type and p-type semiconductors connected in series as shown in Fig. 1(a). The electrons in the n-type material and the holes in the p-type material all carry heat away from the top metal-semiconductor junctions, which leads to a cooling at the junctions called the Peltier effect. Conversely, if a temperature difference is maintained between the two ends of the materials as shown in Fig. 1(b), higher thermal energy electrons and holes will diffuse to the cold side, creating a potential difference that can be used to power an external load. This Seebeck effect is the principle for thermocouples. For each material, the cooling effect is gauged by the Peltier coefficient Π that relates the heat carried by the charges to the electrical current through $Q=\Pi\times I$. The power generation is measured by the Seebeck coefficient S, which relates the voltage generated to the temperature difference through $\Delta V=-S\times\Delta T$. The Peltier and the Seebeck coefficients are related through the Kelvin relation $\Pi=ST$. Practical devices are made of multiple pairs of p-type and n-type semiconductors as shown in Fig. 1(c). Analysis shows that efficient coolers and power generators should have a large figure-of-merit [1],

$$Z=\frac{\sigma S^2}{k},$$

where σ is the electrical conductivity and k the thermal conductivity. The reason that the electrical conductivity σ enters Z is due to the Joule heating in the element. Naturally, the Joule heat should be minimized by increasing the electrical conductivity. The thermal conductivity k appears in the denominator of Z because the thermoelectric elements also act as the thermal insulation between the hot and the cold sides. A high thermal conductivity causes too much heat leakage through heat conduction. Because Z has a unit of inverse temperature, the nondimensional figure of merit ZT is often used. The best ZT materials are found in semiconductors [2]. Insulators have poor electrical conductivity. Metals have relatively low Seebeck coefficients. In addition, the thermal conductivity of a metal, which is dominated by electrons, is proportional to the electrical conductivity, as dictated by the Wiedemann-Franz law. It is thus hard to realize high ZT in metals. In semiconductors, the thermal conductivity consists of contributions from electrons (k_e) and phonons (k_p), with the majority contribution coming from phonons. The phonon thermal conductivity can be reduced without causing too much reduction in the electrical conductivity. A proven approach to reduce the phonon thermal conductivity is through alloying [3]. The mass difference scattering in an alloy reduces the lattice thermal conductivity significantly without much degradation to the electrical conductivity. The traditional cooling materials are alloys of Bi_2Te_3 with Sb_2Te_3 (such as $Bi_{0.5}Sb_{1.5}Te_3$, p-type) and Bi_2Te_3 with Bi_2Se_3 (such as $Bi_2Te_{2.7}Se_{0.3}$, n-type), with a ZT at room temperature approximately equals to one [2]. A typical power generation material is the alloy of silicon and germanium, with a $ZT\sim0.6$ at 700°C. Figure 2 plots the theoretical coefficient of performance (COP) and efficiency of thermoelectric coolers and power generators for different ZT values. Also marked in the figures for comparison are other cooling and power generation technologies. Materials with $ZT\sim1$ are not competitive against the conventional fluid-based cooling and power generation technologies. Thus, solid-state cooler and power generators have only found applications in niche areas, such as cooling of semiconductor lasers and power generation for deep space exploration, although the application areas have been steadily increasing.

While the search for high ZT materials before 1990s have been mostly limited to bulk materials, there has been extensive research in the area of artificial semiconductor structures in the last 30 years for electronic and optoelectronic applications. Various means of producing ultrathin and high quality crystalline layers (such as molecular beam epitaxy and metalorganic chemical vapor deposition) have been used to alter the "bulk" characteristics of the materials. Drastic changes are produced by changing the crystal periodicity (by e.g., depositing alternating layers of different crystals), or by changing the electron dimensionality (by confining the carriers in a plane (quantum well) or in a line (quantum wire, etc.)). Even though electrical and optical properties of these artificial crystalline structures have been extensively studied, much less attention has been paid to their thermal and thermoelectric properties. Thermoelectric properties of low-dimensional structures started to attract attention in the 1990s, in parallel to renewed interests in certain bulk thermoelectric materials such as skutterudites [4]. Compared to the research in bulk materials that emphasizes reducing the thermal conductivity, nanostructures offer the chance of improving both the electron and phonon transport through the use of quantum and classical size and interface effects. Several directions have been explored such as quantum size effects for electrons [5,6], thermionic emission at interfaces [7,8], and interface scattering of phonons [9,10]. Impressive ZT values have been reported in some low-dimensional structures [11,12]. Comprehensive reviews on the progress of thermoelectric

Contributed by the Heat Transfer Division for publication in the JOURNAL OF HEAT TRANSFER. Manuscript received by the Heat Transfer Division July 24, 2001; revision received November 20, 2001. Editor: V. K. Dhir.

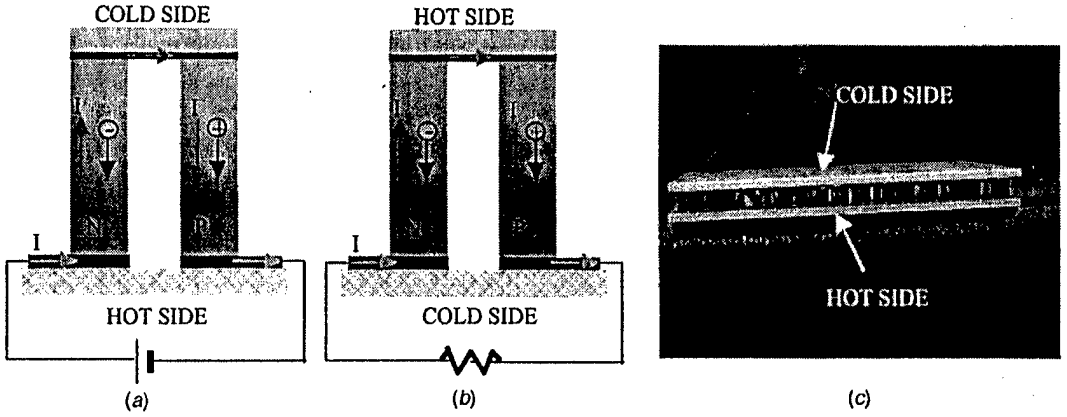


Fig. 1 Illustration of thermoelectric devices (*a*) cooler, (*b*) power generator, and (*c*) an actual device

materials research is presented in a recently published series [4] and in the proceedings of the various international conferences on thermoelectrics held in recent years.

In this article, we have in mind readers interested in nano- and microscale heat transfer and energy conversion and focus on thermoelectric energy conversion in low-dimensional structures. The emphasis is placed on fundamental issues that distinguish energy transport and conversion between nanoscale and macroscale, as well as heat transfer issues related to device development and property characterization. One of our aims is to provide the readers with an overview of recent developments. Because of the wide scope of work being carried out, the cited references are far from complete. Along with the review, we hope to stimulate the readers by pointing out unsolved, challenging questions related to the theory, characterization, and device development.

2 Formulation of Thermoelectric Effects

In solid-state coolers or power generators, heat is carried by charges from one place to another. The current density and heat flux carried by electrons can be expressed as [13]

$$\mathbf{J}(\mathbf{r}) = \frac{1}{4\pi^3}\int\int\int q\mathbf{v}(\mathbf{k})f(\mathbf{r},\mathbf{k})d^3\mathbf{k} \tag{1}$$

$$\mathbf{J}_Q(\mathbf{r}) = \frac{1}{4\pi^3}\int\int\int [E(\mathbf{k})-E_f(\mathbf{r})]\mathbf{v}(\mathbf{k})f(\mathbf{r},\mathbf{k})d^3\mathbf{k}, \tag{2}$$

where q is the unit charge of each carrier, E_f the Fermi energy, $\mathbf{v}$ the carrier velocity, and the integration is over all the possible wavevectors $\mathbf{k}$ of all the charges. The carrier probability distribution function, $f(\mathbf{r},\mathbf{k})$ is governed by the Boltzmann equation. Considering transport processes occurring much slower than the relaxation process and employing the relaxation time approximation, the Boltzmann equation can be expressed as

$$\mathbf{v}\cdot\nabla_r f + \frac{q\boldsymbol{\varepsilon}}{\hbar}\cdot\nabla_{\mathbf{k}} f = -\frac{f(\mathbf{r},\mathbf{k})-f_{eq}(\mathbf{r},\mathbf{k})}{\tau(\mathbf{k})}, \tag{3}$$

where is $\boldsymbol{\varepsilon}$ the electric field, $\tau(\mathbf{k})$ the momentum-dependent relaxation time, $\hbar$ the Planck constant divided by 2π, and f_{eq} the equilibrium distribution function. For electrons and holes,

$$f_{eq}(\mathbf{r},\mathbf{k}) = \frac{1}{1+\exp\left(\dfrac{E(\mathbf{k})-E_f(\mathbf{r})}{k_B T(\mathbf{r})}\right)}, \tag{4}$$

where k_B is the Boltzmann constant, and T the local temperature. Under the further assumption that the local deviation from equilibrium is small, the Boltzmann equation can be linearized and its solution expressed as

$$f(\mathbf{r},\mathbf{k}) = f_{eq}(\mathbf{r},\mathbf{k}) + \tau(\mathbf{k})\mathbf{v}\left(-\frac{\partial f_{eq}}{\partial E}\right)\cdot\left[-\frac{E(k)-E_f}{T}\nabla_r T + q\left(\boldsymbol{\varepsilon}-\frac{1}{q}\nabla_r E_f\right)\right]. \tag{5}$$

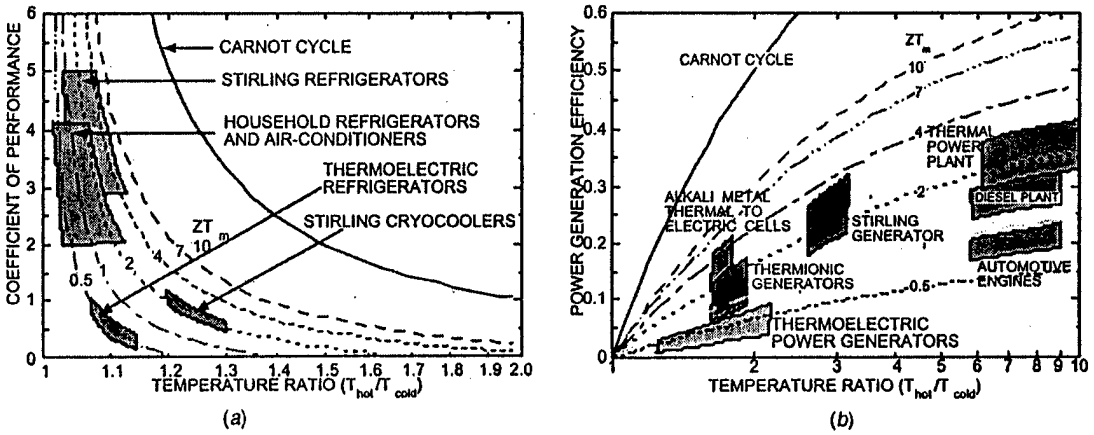


Fig. 2 Comparison of thermoelectric technology with other energy conversion methods for (*a*) cooling and (*b*) power generation

Substituting the above expression in Eqs. (1) and (2) leads to the

$$\mathbf{J}(\mathbf{r}) = q^2 L_0 \left(-\frac{1}{q}\nabla\Phi \right) + \frac{q}{T} L_1 (-\nabla T) \tag{6}$$

$$\mathbf{J}_Q(\mathbf{r}) = q L_1 \left(-\frac{1}{q}\nabla\Phi \right) + \frac{1}{T} L_2 (-\nabla_x T), \tag{7}$$

where Φ is the electrochemical potential $(-\nabla\Phi/q = \varepsilon - \nabla E_f/q)$. The transport coefficients L_n are defined by the following integral

$$L_n = \frac{1}{4\pi^3} \int\int\int \tau(\mathbf{k}) v(\mathbf{k}) v(\mathbf{k}) (E(\mathbf{k}) - E_F)^n \left(-\frac{\partial f_{eq}}{\partial E} \right) d^3\mathbf{k}. \tag{8}$$

From the expressions for $\mathbf{J}$ and $\mathbf{J}_Q$, various material parameters such as the electrical conductivity, thermal conductivity due to electrons, and the Seebeck coefficient can be calculated. For simplicity we assume that both the current flow and the temperature gradient are in the x-direction:

$$\sigma = J_x/(-\nabla\Phi/q)|_{\nabla_x T=0} = q^2 L_0 \tag{9}$$

$$S = (-\nabla\Phi/q)/\nabla_x T|_{J_x=0} = \frac{1}{qT} L_0^{-1} L_1 \tag{10}$$

$$k_e = J_{Q_x}/(-\nabla_x T)|_{J_x=0} = \frac{L_2 L_0 - L_1 L_1}{T L_0}. \tag{11}$$

Rewriting the expressions for electrical conductivity and the thermopower in the form of integrals over the electron energy we get

$$\sigma \equiv \int \sigma(E) \left(-\frac{\partial f_{eq}}{\partial E} \right) dE \tag{12}$$

$$S \equiv \frac{k_B}{q} \frac{\int \sigma(E) \frac{(E-E_F)}{k_B T} \left(-\frac{\partial f_{eq}}{\partial E} \right) dE}{\int \sigma(E) \left(-\frac{\partial f_{eq}}{\partial E} \right) dE} \propto \langle E - E_f \rangle, \tag{13}$$

where we introduced the "differential" conductivity,

$$\sigma(E) \equiv q^2 \tau(E) \int\int v_x^2(E, k_y, k_z) dk_y dk_z \cong q^2 \tau(E) \bar{v}_x^2(E) D(E), \tag{14}$$

where $D(E)$ is the density of states. $\sigma(E)$ is a measure of the contribution of electrons with energy E to the total conductivity. The Fermi "window" factor $(-\partial f_{eq}/\partial E)$ is a bell-shape function centered at $E = E_f$, having a width of $\sim k_B T$. At a finite temperature only electrons near the Fermi surface contribute to the conduction process. In this picture, the thermopower is the "average" energy transported by the charge carriers. In order to achieve the best thermoelectric properties, $\sigma(E)$, within the Fermi window, should be as big as possible, and at the same time, as asymmetric as possible with respect to the Fermi energy.

The thermal conductivity of phonons is also often modeled from the Boltzmann equation under the relaxation time approximation,

$$k_p = \frac{1}{3} \sum \int C(\omega) v_p(\omega) \Lambda(\omega) d\omega, \tag{15}$$

where C is the specific heat of phonons at frequency ω, v_p the phonon group velocity, and Λ the phonon mean free path.

The above formulation for the thermoelectric properties leads to the following possibilities to increase ZT and thus the energy conversion efficiency of devices made of nanostructures.

1. Interfaces and boundaries of nanostructures impose constraints on the electron and phonon waves, which lead to a change in their energy states and correspondingly, their density of states and group velocity.
2. The symmetry of the differential conductivity with respect to the Fermi level can be controlled using quantum size effects and classical interface effects (as in thermionic emission).
3. The phonon thermal conductivity can be reduced through interface scattering and through the alteration of the phonon spectrum in low-dimensional structures.

3 Nanostructures for Solid-State Energy Conversion

The transport of electrons and phonons in nanostructures is affected by the presence of the interfaces and surfaces. Since electrons and phonons have both wave and particle characteristics, the transport can fall into two different regimes: totally coherent transport in which electrons or phonons must be treated as waves and totally incoherent transport in which either or both of them can be treated as particles. There is, of course, the intermediate regime where transport is partially coherent—an area that has not been studied extensively. Whether a group of carriers are coherent or incoherent depend on the strength of phase destroying scattering events (such as internal or diffuse interface scattering). In a nanostructure with no phase-destroying scattering events, a monochromatic wave can experience many coherent scatterings while preserving the phase. The coherent superposition of the incoming and scattered waves leads to the formation of new energy bands for electrons and/or phonons. For example, the quantized energy states of electrons in a quantum well are the result of the formation of standing waves inside the structure. The standing wave can be regarded as the superposition of two counter-propagating waves, each experiencing phase preserving reflections at the interfaces. On the other hand, if there is a strong internal scattering (which can be judged from the momentum relaxation time) or if the interface scattering is not phase preserving (such as due to diffuse scattering), no new energy bands form and the energy states of the carriers in such a structure are identical to these in its bulk material. Electron transport in both coherent and incoherent regimes has been considered and potential benefits of nanostructures for the power factor $(S^2\sigma)$ have been studied. Similarly, phonon heat conduction in both regimes has also be considered, although most studies are based on the particle approach. We will divide the discussion into roughly three categories: (1) improving the electronic power factor based on coherent electron states, (2) improving the electronic energy conversion based on interface filtering for incoherent electrons, (3) improving ZT by reducing the phonon thermal conductivity.

3.1 Electron Energy States Engineering. This approach was suggested in a pioneering work by Dresselhaus and coworkers [5,6]. The main idea is that energy states in nanostructures are very different from those in macrostructures due to the quantum size effects on electrons. Figure 3 illustrates, qualitatively, the density of states (DOS) of electrons in bulk materials,

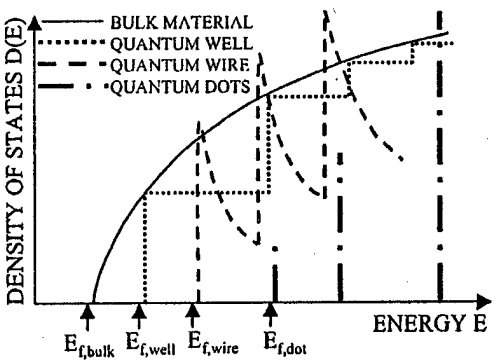


Fig. 3 Schematic illustration of the density-of-states of electrons in bulk, quantum well, quantum wire, and quantum dots materials.

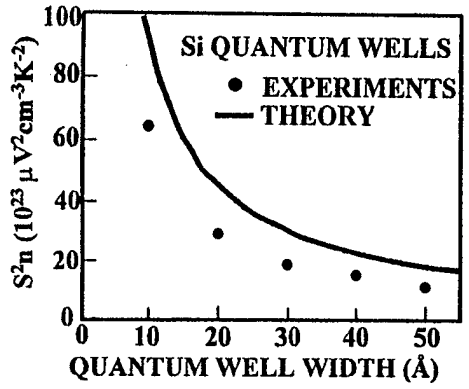


Fig. 4 Product of the Seebeck coefficient square and carrier density as a function of the silicon quantum well width [15]

quantum wells, quantum wires, and quantum dots. Examination of Eq. (13) indicates that the Seebeck coefficient is large when the average electron energy is far apart from the Fermi level. In semiconductors, a large Seebeck coefficient occurs when the Fermi level is inside the band-gap. A Fermi level deep inside the band-gap, however, leads to a low electrical conductivity. The optimized Fermi level usually is close to the band edge. Because the function $\partial f_{eq}/\partial E$ is nonzero only in an energy range $\sim k_B T$ near the Fermi level, the higher the DOS in this range, the larger power factor we can anticipate. In bulk materials, the parabolic shape of the DOS means that the electron density surrounding the Fermi level is small. In quantum structures, the steps and the spikes in the DOS suggest that $S^2\sigma$ can be increased. In a theoretical study by Mahan and Sofo [14], it was suggested that the best thermoelectric materials will have a spike like DOS. Quantum dots fit ideally into such a picture. A single quantum dot, however, is not of much interest for building into useful thermoelectric devices (but may be of interest to create localized cooling on the nanoscale). Thus the study began with quantum wells (extremely thin films) and quantum wires (extremely small wires). Experimental results for transport inside PbTe and Si/SiGe quantum well systems indicated an increase of ZT inside the quantum well, as shown in Fig. 4 [15].

A single quantum well, however, cannot be used to build useful devices because the film is too thin (typically less than a few hundreds angstroms). Multilayer structures were therefore used in the proof-of-concept experiments. For multilayer structures such as superlattices, three questions were raised on the effectiveness of the quantum confinement approach [16,17]. One is that electrons will tunnel through the barrier layer when the barriers are very thin. The second argument is that the barrier does not contribute to the thermoelectric transport but does contribute to the reverse heat conduction. And finally, there is also concern of interface scattering of electrons in narrow wells. A possible approach for improving performance is to utilize the quantum confinement effects inside both the quantum well and the barrier layer, and to have electrons in different carrier pockets in momentum space confined in different regions [18,19]. Additionally, the thermal conductivity of very thin superlattices can be reduced due to interface scattering—a topic we will discuss later on [20]. A natural extension of the quantum well and superlattice theory is to quantum wires. Theoretical studies predict a large enhancement of ZT inside quantum wires. Experimentally, different quantum wire deposition methods have been explored [6].

The experimental results that are inspired by the theoretical studies have proven to be impressive and unexpected [21]. After the proof-of-concept demonstration of ZT enhancement in two-dimensional quantum wells, Harman's group showed that quantum-dot superlattices have a significantly higher power factor than their corresponding bulk materials [22]. Using a thermal conductivity value estimated from bulk properties, ZT values as high as two have been reported. At this stage, no models exist to quantitatively explain the observed increase in ZT.

The above-discussed approaches are based on transport perpendicular to the confinement directions, i.e., along the film plane or wire axis. There are also considerations of the DOS change for electron transport perpendicular to the film plane of the superlattices [23,24]. These calculations, however, do not show a significant increase of the electronic power factor along these directions and suggest that the thermal conductivity reduction may be a more beneficial factor to explore along this transport direction.

Research for improving ZT using quantum confinement of electrons raises several interesting questions related to heat transfer [20]. First, the thermal conductivity along the film plane and wire axis should be reduced due to phonon interface scattering. We will discuss more about this point later. Second, the parasitic thermal conductivity in the barrier layer is still a concern for certain quantum structures that do not utilize the barrier region for electron transport. For example, nanowires can be deposited in an anodized alumina matrix with nanometer scale channels. Although the thermal conductivity of the anodized alumina is relatively small, it is not negligible, and further reduction of the parasitic heat conduction path through such a matrix should be considered [25]. Another very interesting yet little explored aspect is the transport processes in the synthesis of nanowires. Several different methods have been explored, including pressure injection of molten bismuth into the template [26], physical vapor deposition [27], and electrodeposition [28]. Heat and mass transport inside these nanoscale channels could be very different than that in bulk channels. So far, these techniques are developed only through trial-and-error. It is generally found that smaller diameter channels are more difficult to fill, leading to partial filling or discontinuous wires. Systematic studies of the transport processes will help with the optimization of the deposition conditions.

3.2 Heterostructure Integrated Thermionic Refrigeration. Thermionic energy conversion is based on the idea that a high work function cathode in contact with a heat source will emit electrons [29]. These electrons are absorbed by a cold, low work function anode, and they can flow back to the cathode through an external load where they do useful work. Practical vacuum thermionic generators are limited by the work function of available metals or other materials that are used for cathodes. Another important limitation is the space charge effect. The presence of charged electrons in the space between the cathode and anode will create an extra potential barrier between the cathode and anode, which reduces the thermionic current. The materials currently used for cathodes have work functions >0.7 eV, which limits the generator applications to high temperatures >500 K. Mahan [30] proposed these vacuum diodes for thermionic refrigeration. Basically, the same vacuum diodes which are used for generators will work as a cooler on the cathode side and a heat pump on the anode side under an applied bias. Mahan predicted efficiencies of over 80 percent of the Carnot value, but still these refrigerators only work at high temperatures (>500 K). In the following, we will see that heterostructures have a potential to achieve thermionic refrigeration at room temperature [7,31–33]. Vacuum thermionic refrigeration based on resonant tunneling through triangular wells and small gaps in vacuum has been proposed recently [34,35]. Theoretical calculations predict operation at room temperature and below with a cooling power density of 100 W/cm^2. Net cooling based on such vacuum thermionic coolers has yet to be confirmed experimentally.

Using various material systems one can produce different barrier heights in the anode and in the cathode (typically 0 to 0.4 eV). This is determined by the band-edge discontinuity between heterolayers. The heterostructure integrated thermionic coolers (HIT) in Fig. 5 could operate in two modes. In the nonlinear regime, electron transport is dominated by the supply of electrons in the cathode layer. Since only hot electrons (with energy greater than

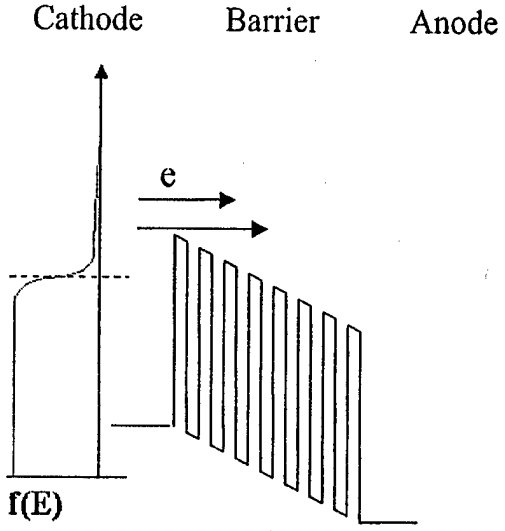
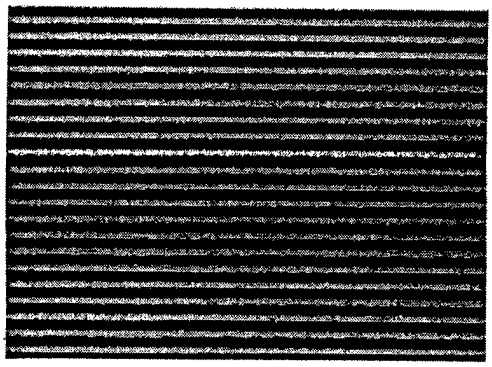


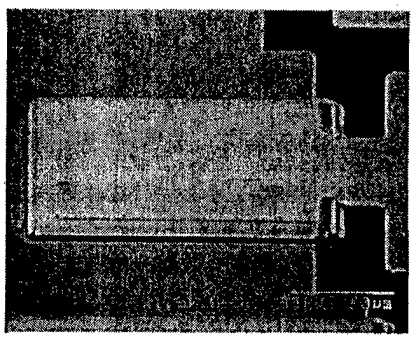
Fig. 5 Heterostructure thermionic emission for cooling at room temperatures.

E_f) are emitted above the barrier, electron-electron and electron-phonon interactions try to restore the quasi Fermi distributions in the cathode layer by absorbing heat from the lattice, thus cooling the layer. This heat is deposited on the anode side. Theoretical estimates by Shakouri and co-workers [7,32] show that there is an optimal barrier width of the order of a few electron energy relaxation lengths and an optimum barrier height of the order of k_BT, and that such heterostructure coolers can provide 20–30°C cooling with KW/cm^2 cooling density. Since the operating currents for the device is very high (10^5 A/cm^2), non-ideal effects such as the Joule heating at the metal-semiconductor contact resistance, and the reverse heat conduction have limited the experimental cooling results to <1°C. There is another regime of operation in which electron transport is dominated by the barrier structure. A superlattice is chosen so that hot electrons move easily in the materials, but the movement of cold electrons is more restricted. In this case, there will be also net cooling in the cathode layer and heating in the anode layer. Shakouri et al. [36] noted that a small barrier height on the order of k_BT does not give much improvement over bulk thermoelectric materials, and suggested tall barriers and high doping densities to achieve a large number of electrons moving in the material. To have a good HIT cooler, the barrier material should simply have an adequate electrical conductivity and a low thermal conductivity, making ternary and quaternary semiconductors good candidates.

On the experimental side, Shakouri and co-workers have fabricated thin-film thermoelectric coolers based on single heterojunction structures [37] and superlattice structures [38–42]. The SiGe/Si superlattice micro coolers can be monolithically integrated with Si-based microelectronic devices to achieve localized cooling and temperature control. Cooling by as much as 4.2 K at 25°C and 12 K at 200°C was measured on 3 μm thick, 60 $\times 60\ \mu$m^2 devices. This corresponds to maximum cooling power densities approaching kW/cm^2. The micro cooler structure is based on cross-plane electrical transport. The main part of the cooler is a 3 μm strain-compensated SiGe/Si superlattice. It consists of 200 periods of (12 nm $Si_{0.75}Ge_{0.25}$/3 nm Si), doped with boron to about 6×10^{19} cm^{-3}. The $Si_{0.75}Ge_{0.25}$/Si superlattice has a valence band offset of about 0.16 eV, and hot holes going over this barrier can produce thermionic cooling. This superlattice was grown using molecular beam epitaxy (MBE). Its average lattice constant is that of $Si_{0.8}Ge_{0.2}$, and a buffer layer is required for it to be grown on a Si substrate. To reduce the material growth time in the MBE system, the buffer layer was grown on a p^+ (001) Si



(*a*)



(*b*)

Fig. 6 (*a*) TEM image of the SiGe/Si superlattice (the dark parts are the 12 nm $Si_{0.75}Ge_{0.25}$ layers, the light parts are the 3 nm Si layers), and (*b*) a scanning electron micrograph of a fabricated micro refrigerator

substrate by chemical vapor deposition (CVD) in the form of a graded SiGe structure. The boron doping in the buffer layer is 5×10^{19} cm^{-3}. Following the superlattice growth, a 0.3 μm $Si_{0.8}Ge_{0.2}$ cap layer was grown with a boron doping level of 2×10^{20} cm^{-3} to get a good ohmic contact to the device. Figure 6 shows a cross-section transmission electron microscopy (TEM) image of the MBE-grown SiGe/Si superlattice, and also a scanning electron micrograph of a micro cooler device. Figure 7 displays the measured cooling on $60\times 60\ \mu$m^2 superlattice cooler and a Si cooler at the heat sink temperature of 25°C. Figure 8 shows the temperature profile at a current of ~400 mA. One notices localized and uniform cooling on top of the micro refrig-

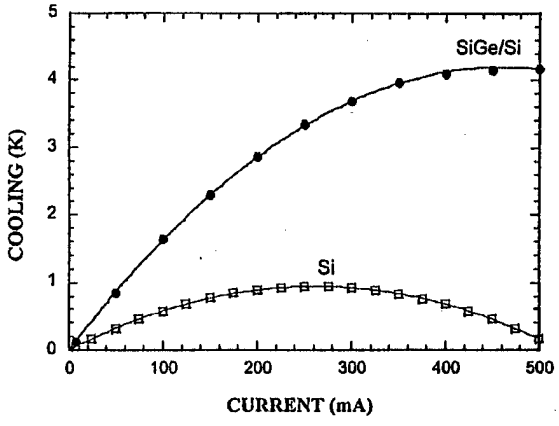


Fig. 7 Cooling measured on $60\times 60\ \mu$m^2 SiGe/Si superlattice coolers and on Si coolers at the heat sink temperature of 25°C

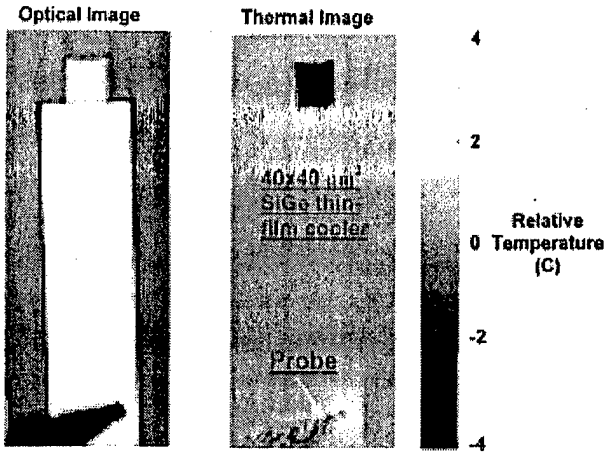


Fig. 8 Temperature distribution on top of a 40×40 micron square SiGe thin film cooler measured using thermoreflectance imaging. The applied current is ~400 mA.

erator as well as Joule heating near the probe on the side contact. With the use of a resistive heat load on top of the micro refrigerator, cooling power densities exceeding 500 W/cm^2 have been demonstrated [43].

Electron and phonon transport perpendicular to interfaces raise interesting heat transfer and energy conversion issues. One example is where heat is generated. Joule heating is often treated as a uniform volumetric heat generation. In heterostructures, the energy relaxation from electrons to phonons occurs over a distance comparable to the film thickness, and heat generation is no longer uniform. For single layer devices, this could benefit the device efficiency in principle [32,44]. Such non-uniform heat generation is a type of hot electron effect that has been studied in electronics [45], and has also be discussed quite extensively in the literature in the context of ultrafast laser-matter interactions [46]. Another example is the concurrent consideration of ballistic electron transport and ballistic phonon transport, coupled with nonequilibrium electron-phonon interaction. Zeng and Chen [47] started from the Boltzmann equations for electrons and phonons and obtained approximate solutions for the electron and phonon temperature distributions in heterostructures, as shown in Fig. 9. In this case, both electron and phonon temperatures show a discontinuity at the interface. The phonon temperature discontinuity is the familiar thermal boundary resistance phenomenon. Zeng and Chen concluded that in the nonlinear transport regime, it is the electron temperature discontinuity at the interface that determines the thermionic effect and the electron temperature gradient inside the film that determines the thermoelectric effect. Similar calculations by Vashaee and Shakouri [48] showed the importance of the electron-phonon coupling coefficient in the optimization of HIT coolers.

3.3 Phonon Thermal Conductivity Reduction Approach. Although phonons do not contribute directly to the energy conversion, the reduction of their contribution to the thermal conductivity is a central issue in thermoelectrics research. Several significant increases in the ZT of bulk materials were due to the introduction of thermal conductivity reduction strategies, such as the alloying [3] and phonon rattler concepts [49]. Size effects on phonon transport have long been established since the pioneering work by Casimir [50] at low-temperatures. Since the 1980s, the thermal conductivity reduction in thin films has drawn increasing attention. Naturally, the phonon thermal conductivity reduction in nanostructures has been considered as beneficial and even as a dominant approach to enhance ZT values.

One proposed approach is to use the thermal conductivity in the direction perpendicular to the superlattice film plane, or the cross-plane direction, while maintaining a low electronic band-edge offset, ideally no offset at all [9]. This would allow the electron transport across the interfaces without much scattering, while phonons would be scattered at the interfaces [51]. Some early experimental data [52,53] indicate that the thermal conductivity of superlattices could be significantly reduced, especially in the cross-plane direction. Tien and Chen [54] suggested the possibility of making super thermal insulators out of superlattices. Extensive experimental data on the thermal conductivity of various superlattices have been reported in recent years [51–65], mostly in the cross-plane direction. Following such a strategy, Venkatasubramanian's group has reported Bi_2Te_3/Se_2Te_3 superlattices with ZT values between 2-3 at room temperature. Although these results need to be confirmed due to the difficulty with the measurements, the reported data so far seem to support such claims and demonstrate that the thermal conductivity reduction is a very effective approach [9].

The mechanisms responsible for thermal conductivity reduction in low-dimensional structures thus have become a topic of considerable debate over the last few years. There have been many studies of the phonon spectrum and transport in superlattices since the original work by Narayanamurti et al. [66], but these works were focused on the phonon modes rather than on heat conduction. The first theoretical modeling predicted a small reduction of

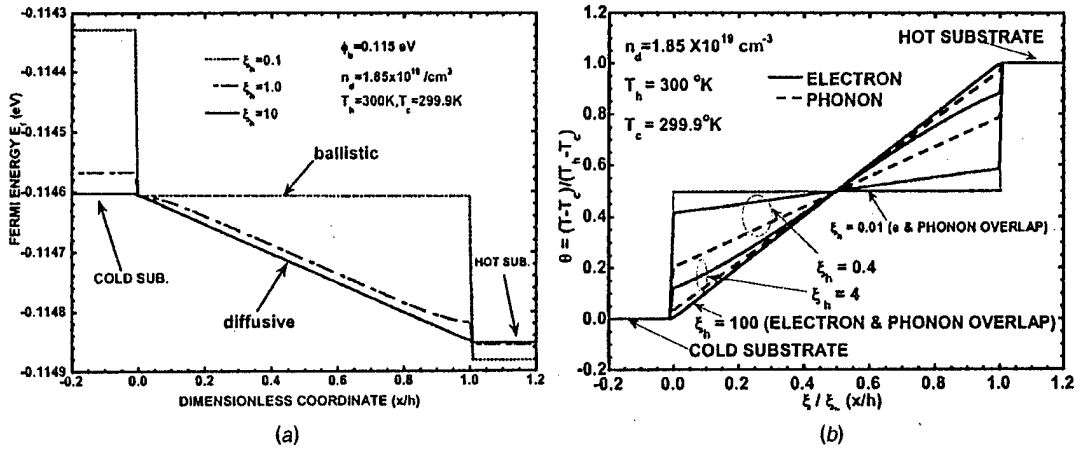


Fig. 9 Distribution of (*a*) Fermi level and (*b*) electron and phonon temperature inside double heterojunction structures. The dimensionless coordinate is normalized to the film thickness. ξ_h is the electron or phonon mean free path divided by the film thickness, n_d the carrier concentration and ϕ_b the barrier height.

the superlattice thermal conductivity [67] due to the formation of minigaps or stop bands. This predicted reduction, however, was too small compared to experiment results in recent years. Two major theoretical approaches were developed in the 1990s to explain the experimental results. One is based on solving the Boltzmann equation with the interfaces of the superlattice treated as boundary conditions [68–71]. The other is based on lattice dynamics calculation of the phonon spectrum and the corresponding change in the phonon group velocity [72–76]. More recently, there are also efforts to use molecular dynamics to simulate the thermal conductivity of superlattices directly [77,78].

Similar to the electron transport in superlattice structures, there could be several different regimes of phonon transport: the totally coherent regime, the totally incoherent regime, and the partially coherent regime. The lattice dynamics lies in the totally coherence regime. Such approaches are based on the harmonic force interaction assumption and thus do not consider anharmonic effects. A bulk relaxation time is often assumed. The main results out of the lattice dynamics models is that the phonon group velocity reduction caused by the spectrum change can lower the thermal conductivity by a factor of ~7–10 at room temperature for Si/Ge superlattices, and by a factor of 3 for GaAs/AlAs superlattices. Although it can be claimed that the predicted reduction in Si/Ge system is of the order of magnitude that is experimentally observed, the prediction clearly cannot explain the experimental results for GaAs/AlAs superlattices. The lattice dynamics model also showed that when the layers are 1–3 atomic-layers thick, there is a recovery of the thermal conductivity. The acoustic wave based model [79], which treats the superlattices as an inhomogeneous medium, shows a similar trend. It reveals that the thermal conductivity recovery is due to phonon tunneling and that the major source of the computed thermal conductivity reduction in the lattice dynamics model is the total internal reflection, which in the phonon spectrum representation, causes a group velocity reduction. For experimental results so far, the explanation of the thermal conductivity reduction based on the group velocity reduction has not been satisfactory even for the cross-plane direction. For the in-plane direction, the group velocity reduction alone leads to only a small reduction in thermal conductivity [76], and cannot explain the experimental data on GaAs/AlAs and Si/Ge superlattices [52,55,65]. There is a possibility that the change in the phonon spectrum creates a change in the relaxation time [80] but such a mechanism is unlikely to explain the experimental results for relatively thick-period superlattices since the density of states does not change in these structures [76].

Boltzmann equation-based models that treat phonons as particles transporting heat in inhomogeneous layers lie in the totally incoherent regime [68,71]. Theoretical calculations have been able to explain quantitatively the experimental data. The models are based on the solution of the Boltzmann equation using the relaxation time in the bulk materials for each layer. Phonon reflection and transmission at the interfaces are modeled based on past studies of the thermal boundary resistance. Compared to the lattice dynamics and acoustic waves models, the particle model allows the incorporation of diffuse interface scattering of phonons. In the models presented so far, the contribution of diffuse scattering has been left as a fitting parameter. In Fig. 10, we show the experimental in-plane and cross-plane thermal conductivity of a Si/Ge superlattice, together with simulation results based on the Boltzmann equation. One argument for the validity of the particle model is that thermal phonons have a short thermal wavelength, which is a measure of the coherence properties of broadband phonons inside the solid [68]. It is more likely, however, that the diffuse interface scattering, if it indeed happens as models suggested, destroys the coherence of monochromatic phonons and thus prevents the formation the superlattice phonon modes. The particle-based modeling can capture the effects of total internal reflection, which is partially responsible for the large group velocity reduction under the lattice dynamics models. Approximate

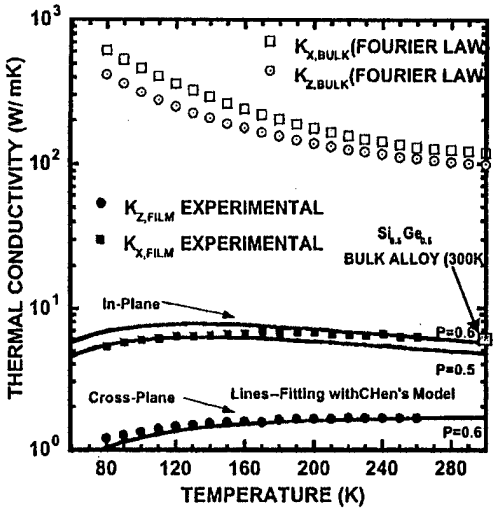


Fig. 10 Anisotropic thermal conductivity of the strained Si/Ge (20 Å /20 Å) superlattice: experimental data were fitted using Chen's models [68,71]. Also shown in the figure are comparisons of experimental data experimental data with predictions of Fourier theory based on bulk properties of each layer, and with compositionally equivalent alloy (300K) [65].

methods to incorporate phonon confinement or inelastic boundary scattering are also proposed. From the existing modeling, it can be concluded that for heat flow parallel to the interfaces, diffuse interface scattering is the key factor causing the thermal conductivity reduction. For the case of heat conduction perpendicular to the interfaces, phonon reflection, confinement, as well as diffuse scattering can greatly reduce the heat transfer and thermal conductivity. The larger the reflection coefficient, the larger is the thermal conductivity reduction in the cross-plane direction.

A key unsolved issue is what are the actual mechanisms of phonon scattering at the interfaces, particularly what causes the diffuse phonon scattering. Phonon scattering has been studied quite extensively in the past in the context of thermal boundary resistance. Superlattice structures that are grown by epitaxy techniques usually have better interface morphology than the other types of interfaces studied previously. Even for the best material system such as GaAs/AlAs, however, the interfaces are not perfect. There is interface mixing and there are also regions with monolayer thickness variation. These are naturally considered as potential sources of diffuse interface scattering for which a simplified model was developed by Ziman [82]. Another possibility is the anharmonic force between the atoms in two adjacent layers. Boltzmann equation-based modeling assumes a constant parameter p to represent the fraction of phonons specularly scattered. Ju and Goodson [81] used an approximate frequency-dependent expression for p given by Ziman [82] in the interpretation of the thermal conductivity of single layer silicon films. Chen [71] also argued that inelastic scattering occurring at interfaces can provide a path for the escape of confined phonons. A promising approach to resolve this issue is molecular dynamics simulation [77,78]. In addition to the interface scattering mechanisms, there are also several other unanswered questions. For example, experimental data of Venkatasubrmanian seems to indicate a butterfly-shaped thermal conductivity curve as a function of thickness [9,83]. Quantitative modeling of the stress and dislocation effects also needs to be further refined.

Since the lattice dynamics and the particle models present the totally coherent and totally incoherent regimes, a theoretical approach that can include both effects should be sought. Simkin and Mahan [84] proposed a new lattice dynamics model by the introduction of an imaginary wavevector that is related to the mean free path. This approach leads to the prediction of a minimum in

the thermal conductivity value as a function of the superlattice period thickness. For thicknesses larger than the minimum, the thermal conductivity increases with thickness and eventually approaches the bulk values. For thicknesses thinner than the minimum, the thermal conductivity recovers to a higher value. However, it should be pointed out that the imaginary wavevector represents an absorption process, not exactly a scattering process, as is clear in the meaning of the extinction coefficient of the optical constants. It will be interesting to see whether such an approach can explain the experimental observed trends of thermal conductivity reduction along the in-plane direction.

Aside from superlattices and thin films, other low-dimensional structures such as quantum wires and quantum dots are also being considered for thermoelectric applications. There are a few experimental and theoretical studies on the thermal conductivity of quantum dot arrays and nanostructured porous medium [85,86]. Theoretically, one can expect a larger thermal conductivity reduction in quantum wires compared to thin films [87,88]. The measurements of the thermal conductivity in quantum wires have been challenging. Recent measurements on the thermal conductivity of carbon nanotubes provide possible approach for measurements on nanowires for thermoelectric applications [89]. Nanowires for thermoelectric applications, however, usually have a low thermal conductivity, which may need different characterization techniques.

4 Characterization

The characterization of thermoelectric properties has turned out to be the most challenging issue for the development of nanostructure-based thermoelectric materials. First, the thermal conductivity measurements for even bulk materials are not easy. For thin films, these measurements become even more difficult. Even the normally easier measurements in bulk materials, such as for the electrical conductivity and the Seebeck coefficient, can be complicated due to the small thickness of the film and the contributions from the substrate.

It is generally recognized that the thermal conductivity is the most difficult parameter to measure. Fortunately, thin film thermal conductivity measurements have drawn considerable attention over the past two decades and different methods have been developed. One popular method to measure the thermal conductivity of thin films is the 3ω method [90,91]. For thermoelectric thin films such as superlattices, there are several complications. For example, thermoelectric films are semiconductors and thus an insulating film is required between the heater and the film. The superlattice thermal conductivity is highly anisotropic. The 3ω method is typically applied to measure the cross-plane thermal conductivity by ensuring that the heater width is much larger than the film thickness. Often, an additional buffer layer exists between the film and the substrate. For Si/Ge, the buffer is graded and thus has a continuously varying thermal conductivity profile. In applying the 3ω method, there is also the contrast factor that must be considered between the film and the substrate. When the film and the substrate have close properties, more complicated modeling is needed. By careful modeling and experimental design, the 3ω method can be applied to a wide range of thin films for measuring the thermal conductivity in both in-plane and cross-plane directions [92]. Other methods such as ac calorimetry, photothermal and pump-and-probe methods have also been used to measure the thermal diffusivity of superlattices. References [10], [93–96] provide detailed reviews of existing methods. By assuming that the specific heat does not change much, which is usually a valid assumption, the thermal conductivity of the structures can be calculated from the measured diffusivity.

Although the measurement of the electrical conductivity and Seebeck coefficient is considered relatively straight-forward for bulk materials, it has turned out to be much more complicated for thin films. For transport along the thin film plane, the complications arise from the fact that most thin films are deposited on semiconductor substrates and the thermoelectric effect of the substrates can overwhelm that of the films. To circumvent these difficulties, several approaches have been taken, such as removing the substrate or growing the film on insulating layers. For example, Si/Ge superlattices are grown on silicon-on-insulator structures. Even with these precautions, there are still complications such as the existence of the buffer. Thus, differential measurements are sometimes used to subtract the influence of the buffer layer. For transport in the cross-plane direction, measurements of the electrical conductivity and Seebeck effect become much more difficult because the films are usually very thin. The recently reported ZT values between 2–3 for Bi_2Te_3/Se_2Te_3 superlattices were obtained using the transient Harman method [9]. Although the method is well established for bulk materials, the application to thin film structures requires careful consideration of various heat losses and heat generation through the leads. This is a topic that has yet to be fully addressed. In addition, the transient Harman method gives ZT rather than individual thermoelectric properties such as the Seebeck coefficient. A comparative method was recently developed to measure the Seebeck coefficient in the cross-plane direction of the superlattices [97].

5 Micro Devices and Potential Applications

Although thermoelectric coolers and power generators are intrinsically smaller than conventional coolers and power generators, on the order of a few millimeters to centimeters, there are major efforts to develop microscale thermoelectric devices for several reasons. (1) Monolithic integration is desirable for the cooling and temperature stabilization of electronic and phononic devices. (2) Heat flux that can be handled by coolers increases as the device thickness decreases. Of course, smaller device sizes are attractive for their weight and volume. Several types of thermoelectric microdevices are being developed. Figures 7 and 8 show device performance based on Si/Ge superlattices with transport in the cross-plane direction. Another approach is to use electrodeposited films [98] to make microdevices. Efforts exist in using micromachining of bulk thermoelectric materials to make microdevices [99,100]. Devices based on transport along the film plane are also being pursued, for applications in detectors and power sources [101,102].

As the device length becomes smaller, several degradation factors become important and must be addressed. These include (1) the electrical contact resistance, (2) thermal contact resistance, (3) heat sink thermal resistance in both the hot and the cold sides, and (4) additional heat leaks caused by contacts. For transport perpendicular to the film plane, the most important issues are the electrical contact resistance, the thermal boundary resistance and the spreading resistance in the hot and the cold regions. For transport along the film plane, heat leakage through the support layers normally requires the removal of the substrate. Even without a substrate, the heat leakage through the supporting membrane and the buffer layers can significantly degrade the device performance, particularly when the thermoelectric film is very thin.

5.1 Applications. Thermoelectric microdevices have some immediate applications. If the reported ZT is further confirmed and enhanced, the applications will undoubtly expand into many areas. Here, we will discuss a number of potential applications: (1) temperature stabilization, (2) high cooling density spot cooling, and (3) micropower generation.

Temperature stabilization is very important for optoelectronic devices such as laser sources, switching/routing elements, and detectors. They require careful control over their operating temperature. This is especially true in current high speed and wavelength division multiplexed (WDM) optical communication networks. Long haul optical transmission systems operating around 1.55 μm wavelength typically use erbium-doped fiber amplifiers (EDFA's), and are restricted in the wavelengths they can use due to the finite bandwidth of these amplifiers. As more channels are packed into this wavelength window, the spacing between adjacent channels

becomes smaller and wavelength drift becomes very important. Temperature variations are the primary cause for the wavelength drift, and they also affect the threshold current and output power in laser sources. Most stable sources such as distributed feedback (DFB) lasers and vertical cavity surface emitting lasers (VCSEL's) can generate large heat power densities on the order of kW/cm^2 over areas as small as 100 μm^2 [103,104]. The output power for a typical DFB laser changes by approximately 0.4 dB/°C. Typical temperature-dependent wavelength shifts for these laser sources are on the order of 0.1 nm/°C [105]. Therefore, a temperature change of only a few degrees in a WDM system with a channel spacing of 0.2–0.4 nm would be enough to switch data from one channel to the adjacent one, and even less of a temperature change could dramatically increase the crosstalk between two channels. Temperature stabilization or refrigeration is commonly performed with conventional thermoelectric (TE) coolers. However since their integration with optoelectronic devices is difficult [103,106], component cost is greatly increased because of packaging. The reliability and lifetime of packaged modules are also usually limited by their TE coolers [107]. Microdevices monolithically integrated with the functioning optoelectronic devices have advantages over separate devices in terms of their response time, size, and costs.

Many electronic and optoelectronic devices dissipate high heat flux. Conventional thermoelectric devices cannot handle large heat flux. With reduced leg length, the cooling heat flux of thermoelectric devices increases, thus providing the opportunity to handle high heat flux devices. It should be remembered, however, that more heat flux must be rejected at the hot side and must be removed using conventional heat transfer technologies such as heat pipes and high thermal conductivity heat spreaders. The active cooling method is beneficial only when the device needs to be operated below ambient temperatures or for temperature stabilization. Examples are infrared detectors and quantum cascade lasers. The speed of many electronic devices increases with reduced temperature and thus it is possible to use thermoelectric coolers to gain increased speed. Instead of cooling the whole chip, thermoelectric microcoolers can potentially be applied to handle local hot spots in semiconductor chips [108]. Regions with sizes ranging from 10's to 100's of micron in diameter have a temperature 10–30°C higher than the average chip temperature. This causes clock delays and failures in digital circuits. In addition, chip reliability due to electromigration is a thermally activated process, so the mean-free time between failures decreases exponentially as the temperature rises.

Thermoelectric devices have traditionally been used as radiation detectors such as thermopiles and can be used as power sources. With the rapid developments in MEMS, microscale power supply has been in increasing demands. Thermoelectric micro-generators can be coupled with environmental heat sources to drive sensors and microdevices for automanous operation of these devices. The body temperature powered wristwatch is a recent example [109].

6 Concluding Remarks

In this paper, we discuss recent progress in nanostructure-based solid-state energy conversion. Energy transport for both electrons and phonons can differ significantly from that in bulk materials. The nanoscale size effects can be used to improve the energy conversion efficiency. Recent studies have led to quite a large increase in ZT values and significant new insights into thermoelectric transport in nanostructures. There is, however, much left to be done in new materials syntheses, characterization, physical understanding, and device fabrication. This is a research area that the heat transfer community can both benefit from and contribute to. Meanwhile, we would like to emphasize that thermoelectric materials research is a multidisciplinary endeavor and requires close collaboration among researchers to address issues in materials, theory, characterization, and devices. Among these issues, heat transfer plays a significant role.

Acknowledgments

G. C. would like to acknowledge supports from NSF, DoD/ONR MURI on Thermoelectrics and DARPA HERETIC project for the work reported here. He would like to thank Professors M. S. Dresselhaus, K. L. Wang, T. Sands, and R. Gronsky, and their students for their contributions to the discussed work. A. S. acknowledges support from the Packard Foundation, NSF and DARPA HERETIC project. He would like to thank Professors J. Bowers, A. Majumdar, V. Narayanamurti, their students, and Dr. Ed. Croke for their contributions to the work presented here.

Nomenclature

C = spectral volumetric specific heat, J m^{-3}Hz^{-1}K^{-1}
D = density of states per unit volume, J^{-1}m^{-3}
E = energy relative to band edge, J
E_f = Fermi level relative to band edge, J
f = electron distribution function
f_{eq} = equilibrium distribution function
$\mathbf{J}$ = current density, A m^{-2}
$\mathbf{J}_q$ = heat flux, W m^{-2}
k = thermal conductivity, W m^{-1}K^{-1}
$\mathbf{k}$ = wavevector
k_B = Boltzmann constant, J K^{-1}
L_n = transport coefficients defined by Eq. (8)
q = charge per carrier, C
$\mathbf{r}$ = coordinate
S = Seebeck coefficient, V K^{-1}
t = time, s
T = temperature, K
v = velocity, m s^{-1}
Z = figure of merit, K^{-1}
ε = electrical field, V/m
Φ = electrochemical potential, J
σ = electrical conductivity Ω^{-1}m^{-1}, or differential conductivity, Ω^{-1}m^{-1}J^{-1}
τ = relaxation time, s
ω = angular frequency, rad . Hz

References

[1] Goldsmid, H. J., 1964, *Thermoelectric Refrigeration*, Plenum Press, New York.
[2] Rowe, D. M., ed., 1995, *Handbook of Thermoelectrics*, CRC Press, Boca Raton.
[3] Ioffe, A. F., 1957, *Semiconductor Thermoelements and Thermoelectric Cooling*, Infosearch Limited, London.
[4] Tritt, T. M., ed., 2001, "Recent Trend in Thermoelectric Materials Research," in *Semiconductors and Semimetals*, Vol. 69-Vol. 71, Academic Press, San Diego.
[5] Hicks, L. D., and Dresselhaus, M. S., 1993, "Effect of Quantum-Well Structures on the Thermoelectric Figure of Merit," Phys. Rev. B, **47**, pp. 12727–12731.
[6] Dresselhaus, M. S., Lin, Y. M., Cronin, S. B., Rabin, O., Black, M. R., Dresselhaus, G., and Koga, T., 2001, "Quantum Wells and Quantum Wires for Potential Thermoelectric Applications," in Semicond. Semimetals, **71**, pp. 1–121.
[7] Shakouri, A., and Bowers, J. E., 1997, "Heterostructure Integrated Thermionic Coolers," Appl. Phys. Lett., **71**, pp. 1234–1236.
[8] Mahan, G. D., 2001, "Thermionic Refrigeration," Semicond. Semimetals, **71**, pp. 157–174.
[9] Venkatasubramanian, R., 2001, "Phonon Blocking Electron Transmitting Superlattice Structures as Advanced Thin Film Thermoelectric Materials," Semicond. Semimetals, **71**, pp. 175–201.
[10] Chen, G., 2001, "Phonon Transport in Low-Dimensional Structures," Semicond. Semimetals, **71**, pp. 203–259.
[11] Harman, T. C., Taylor, P. J., Spears, D. L., and Walsh, M. P., 2000, "Thermoelectric Quantum-Dot Superlattices with High ZT," J. Electron. Mater., **29**, pp. L1–L4.
[12] Dubios, L., 1999, "An Introduction to the DARPA Program in Advanced Thermoelectric Materials and Devices," *Proc. Int. Conf. Thermoelectrics*, ICT'99, pp. 1–4.
[13] Aschroft, N. W., and Mermin, N. D., 1976, *Solid State Physics*, Saunders College Publishing, Fort Worth.

[14] Mahan, G. D., and Sofo, J. O., 1996, "The Best Thermoelectric," Proc. Natl. Acad. Sci. U.S.A., **93**, pp. 7436–7439.
[15] Sun, X., Cronin, S. B., Liu, J. L., Wang, K. L., Koga, T., Dresselhaus, M. S., and Chen, G., 1999, "Experimental Study of the Effect of the Quantum Well Structures on the Thermoelectric Figure of Merit in Si/SixGel-x System," *Proceedings of Int. Conf. Thermoelectrics*, ICT'99, pp. 652–655.
[16] Broido, D. A., and Reinecke, T. L., 1995, "Effect of Superlattice Structure on the Thermoelectric Figure of Merit," Phys. Rev. B, **51**, pp. 13797–13800.
[17] Sofo, J. O., and Mahan, G. D., 1994, "Thermoelectric Figure of Merit of Superlattices," Appl. Phys. Lett., **65**, pp. 2690–2692.
[18] Koga, T., Sun, X., Cronin, S. B., and Dresselhaus, M. S., 1998, "Carrier Pocket Engineering to Design Superior Thermoelectric Materials Using GaAs/AlAs Superlattices," Appl. Phys. Lett., **73**, pp. 2950–2952.
[19] Koga, T., Cronin, S. B., Dresselhaus, M. S., Liu, J. L., and Wang, K. L., 2000, "Experimental Proof-of-Principle Investigation of Enhanced $Z_{3D}T$ in (001) Oriented Si/Ge Superlattices," Appl. Phys. Lett., **77**, pp. 1490–1492.
[20] Chen, G., 1997, "Size and Interface Effects on Thermal Conductivity of Superlattices and Periodic Thin-Film Structures," ASME J. Heat Transfer , **119**, pp. 220–229 (see also Proc. 1996 Nat. Heat Transf. Conf., ASME HTD-Vol. 323, **121**, 1996).
[21] Hicks, L. D., Harman, T. C., and Dresselhaus, M. S., 1996, "Experimental Study of the Effect of Quantum-Well Structures on the Thermoelectric Figure of Merit," Phys. Rev. B, **53**, pp. 10493–10496.
[22] Harman, T. C., Taylor, P. J., Spears, D. L., and Walsh, M. P., 2000, "Thermoelectric Quantum-Dot Superlattices With High ZT," J. Electron. Mater., **29**, pp. L1–L4.
[23] Whitlow, L. W., and Hirano, T., 1995, "Superlattice Application to Thermoelectricity," J. Appl. Phys., **78**, pp. 5460–5466.
[24] Radtke, R. J., Ehrenreich, H., and Grein, C. H., 1999, "Multilayer Thermoelectric Refrigeration in $Hg_{1-x}Cd_xTe$ Superlattices," J. Appl. Phys., **86**, pp. 3195–3198.
[25] Kumar, R., Borca-Tasciuc, D., Zeng, T., and Chen, G., 2000, "Thermal Conductivity of Nanochanneled Alumina," *Proc. Int. Mech. Eng. Congress and Exhibition* (IMECE2000), ASME HTD-Vol. 366-2, pp. 393–398.
[26] Zhang, Z. B., Gekhtman, D., Dresselhaus, M. S., and Ying, J. Y., 1999, "Processing and Characterization of Single-Crystalline Ultrafine Bismuth Nanowires," Chem. Mater., **11**, pp. 1659–1665.
[27] Heremans, J., Thrush, C. M., Lin, Y. M., Cronin, S., and Dresselhaus, M. S., 2000, "Bismuth Nanowire Arrays: Synthesis and Galvanomagnetic Properties," Phys. Rev. B, **61**, pp. 2921–2930.
[28] Behnke, J. F., Prieto, A. L., Stacy, A. M., and Sands, T., 1999, "Electrodeposition of $CoSb_3$ Nanowires," ICT'99 pp. 451–453.
[29] Hatsopoulos, G. N., and Kaye, J., 1958, "Measured Thermal Efficiencies of a Diode Configuration of a Thermo Electron Engine," J. Appl. Phys., **29**, pp. 1124–1125.
[30] Mahan, G. D., 1994, "Thermionic Refrigeration," J. Appl. Phys., **76**, pp. 4362–4366.
[31] Mahan, G. D., and Woods, L. M., 1998, "Multilayer Thermionic Refrigeration," Phys. Rev. Lett., **80**, pp. 4016–4019.
[32] Shakouri, A., Lee, E. Y., Smith, D. L., Narayanamurti, V., and Bowers, J. E., 1998, "Thermoelectric Effects in Submicron Heterostructure Barriers," Microscale Thermophys. Eng., **2**, pp. 37–42.
[33] Moyzhes, B., and Nemchinsky, V., 1998, "Thermoelectric Figure of Merit of Metal-Semiconductor Barrier Structure based on Energy Relaxation Length," Appl. Phys. Lett., **73**, pp. 1895–1897.
[34] Miskovsky, N. M., and Cutler, P. H., 1999, "Microelectronic Cooling Using the Nottingham Effect and Internal Field Emission in a Diamond (Wide-Band Gap Material) Thin-Film Device," Appl. Phys. Lett., **75**, pp. 2147–2149.
[35] Korotkov, A. N., and Likharev, K. K., 1999, "Possible Cooling by Resonant Fowler-Nordheim Emission," Appl. Phys. Lett., **75**, pp. 2491–2493.
[36] Shakouri, A., Labounty, C., Abraham, P., Piprek, J., and Bowers, J. E., 1998, "Enhanced Thermionic Emission Cooling in High Barrier Superlattice Heterostructures," Mater. Res. Soc. Symp. Proc., **545**, pp. 449–458.
[37] Shakouri, A., LaBounty, C., Piprek, J., Abraham, P., and Bowers, J. E., 1999, "Thermionic Emission Cooling in Single Barrier Heterostructures," Appl. Phys. Lett., **74**, pp. 88–89.
[38] Zeng, G. H., Shakouri, A., La Bounty, C., Robinson, G., Croke, E., Abraham, P., Fan, X. F., Reese, H., and Bowers, J. E., 1999, "SiGe Micro-Cooler," Electron. Lett., **35**, pp. 2146–2147.
[39] Fan, X. F., Zeng, G. H., LaBounty, C., Bowers, J. E., Croke, E., Ahn, C. C., Huxtable, S., Majumdar, A., and Shakouri, A., 2001, "SiGeC/Si Superlattice Microcoolers," Appl. Phys. Lett., **78**, pp. 1580–1582.
[40] Fan, X. F., Zeng, G., Croke, E., LaBounty, C., Ahn, C. C., Vashaee, D., Shakouri, A., and Bowers, J. E., 2001, "High Cooling Power Density SiGe/Si Micro-Coolers," Electron. Lett., **37**, pp. 126–127.
[41] LaBounty, C., Shakouri, A., Abraham, P., and Bowers, J. E., 2000, "Monolithic Integration of Thin-Film Coolers with Optoelectronic Devices," Opt. Eng., **39**, pp. 2847–2852.
[42] LaBounty, C., Shakouri, A., and Bowers, J. E., 2001, "Design and Characterization of Thin Film Microcoolers," J. Appl. Phys., **89**, pp. 4059–4064.
[43] Zeng, G., Fan, X., LaBounty, C., Bowers, J. E., Croke, E., and Shakouri, A., "Direct Measurements of Cooling Power Density for SiGe/Si Superlattice Microcoolers," submitted for publication.
[44] Zeng, T. F., and Chen, G., 2000, "Energy Conversion in Heterostructures for Thermionic Cooling," Microscale Thermophys. Eng., **4**, pp. 39–50.
[45] Reggiani, L., ed., 1985, *Hot-Electron Transport in Semiconductors*, Springer-Verlag.
[46] Qiu, T. Q., and Tien, C. L., 1993, "Heat Transfer Mechanisms During Short-Pulse Laser Heating of Metals," ASME J. Heat Transfer , **115**, pp. 835–841.
[47] Zeng, T. F., and Chen, G., 2000, "Nonequilibrium Nonequilibrium Electron and Phonon Transport in Heterostructures for Energy Conversion," *Proceedings of Int. Mech. Eng. Congress and Exhibition* (IMECE2000), ASME HTD-Vol. 366-2, pp. 361–372.
[48] Vashaee, D., and Shakouri, A., 2002, unpublished.
[49] Nolas, G. S., Slack, G. A., Morelli, D. T., Tritt, T. M., and Ehrlich, A. C., 1996, "The Effect of Rare-Earth Filling on the Lattice Thermal Conductivity of Skutterudites," J. Appl. Phys., **79**, pp. 4002–4008.
[50] Casimir, H. B. G., 1938, "Note on the Conduction of Heat in Crystals," Physica, **5**, pp. 495–500.
[51] Venkatasubramanian, R., 1996, "Thin-Film Superlattice and Quantum-Well Structures—A New Approach to High-Performance Thermoelectric Materials," Nav. Res. Rev., **58**, pp. 44–54.
[52] Yao, T., 1987, "Thermal Properties of AlAs/GaAs Superlattices," Appl. Phys. Lett., **51**, pp. 1798–1800.
[53] Chen, G., Tien, C. L., Wu, X., and Smith, J. S., 1994, "Measurement of Thermal Diffusivity of GaAs/AlGaAs Thin-Film Structures," ASME J. Heat Transfer , **116**, pp. 325–331.
[54] Tien, C. L., and Chen, G., 1994, "Challenges in Microscale Conductive and Radiative Heat Transfer," ASME J. Heat Transfer , **116**, pp. 799–807.
[55] Yu, X. Y., Chen, G., Verma, A., and Smith, J. S., 1995, "Temperature Dependence of Thermophysical Properties of GaAs/AlAs Periodic Structure," Appl. Phys. Lett. **67**, pp. 3554–3556, p. 1303.
[56] Capsinski, W. S., and Maris, H. J., 1996, "Thermal Conductivity of GaAs/AlAs Superlattices," Physica B, **220**, pp. 699–701.
[57] Lee, S. M., Cahill, D. G., and Venkatasubramanian, R., 1997, "Thermal Conductivity of Si-Ge Superlattices," Appl. Phys. Lett., **70**, pp. 2957–2959.
[58] Yamasaki, I., Yamanaka, R., Mikami, M., Sonobe, H., Mori, Y., and Sasaki, T., 1998, "Thermoelectric Properties of Bi_2Te_3/Sb_2Te_3 Superlattice Structures," *Proc. Int. Conf. Thermoelectrics*, ICT'98, pp. 210–213.
[59] Capinski, W. S., Maris, H. J., Ruf, T., Cardona, M., Ploog, K., and Katzer, D. S., 1999, "Thermal-Conductivity Measurements of GaAs/AlAs Superlattices Using a Picosecond Optical Pump-and-Probe Technique," Phys. Rev. B, **59**, pp. 8105–8113.
[60] Borca-Tasciuc, T., Liu, W. L., Zeng, T., Song, D. W., Moore, C. D., Chen, G., Wang, K. L., Goorsky, M. S., Radetic, T., Gronsky, R., Koga, T., and Dresselhaus, M. S., 2000, "Thermal Conductivity of Symmetrically Strained Si/Ge Superlattices," Superlattices Microstruct., **28**, pp. 119–206.
[61] Song, D. W., Liu, W. L., Zeng, T., Borca-Tasciuc, T., Chen, G., Caylor, C., and Sands, T. D., 2000, "Thermal Conductivity of Skutterudite Thin Films and Superlattices," Appl. Phys. Lett., **77**, pp. 3854–3856.
[62] Venkatasubramanian, R., 2000, "Lattice Thermal Conductivity Reduction and Phonon Localizationlike Behavior in Superlattice Structures," Phys. Rev. B, **61**, pp. 3091–3097.
[63] Huxtable, S. T., Shakouri, A., LaBounty, C., Fan, X., Abraham, P., Chiu, Y. J., Chiu, Y. J., and Majumdar, A., 2000, "Thermal Conductivity of Indium Phosphide-Based Superlattices," Microscale Thermophys. Eng., **4**, pp. 197–203.
[64] Borca-Tasciuc, T., Achimov, D., Liu, W. L., Chen, G., Ren, H.-W., Lin, C.-H., and Pei, S. S., 2001, "Thermal Conductivity of InAs/AlSb Superlattices," Microscale Thermophys. Eng., **5**, pp. 225–231.
[65] Liu, W. L., Borca-Tasciuc, T., Chen, G., Liu, J. L., and Wang, K. L., 2001, "Anisotropy Thermal Conductivity of Ge-Quantum Dot and Symmetrically Strained Si/Ge'Superlattice," J. Nanosci. Nanotech., **1**, pp. 39–42.
[66] Narayanamurti, V., Stormer, J. L., Chin, M. A., Gossard, A. C., and Wiegmann, W., 1979, "Selective Transmission of High-Frequency Phonons by a Superlattice: the Dielectric Phonon Filter," Phys. Rev. Lett., **43**, pp. 2012–2015.
[67] Ren, S. Y., and Dow, J., 1982, "Thermal Conductivity of Superlattices," Phys. Rev. B, **25**, pp. 3750–3755.
[68] Chen, G., 1997, "Size and Interface Effects on Thermal Conductivity of Superlattices and Periodic Thin-Film Structures," ASME J. Heat Transfer, **119**, pp. 220–229.
[69] Hyldgaard, P., and Mahan, G. D., 1996, "Phonon Knudson Flow in Superlattices," Therm. Conduct., **23**, p. 172–181.
[70] Chen, G., and Neagu, M., 1997, "Thermal Conductivity and Heat Transfer in Superlattices," Appl. Phys. Lett., **71**, pp. 2761–2763.
[71] Chen, G., 1998, "Thermal Conductivity and Ballistic Phonon Transport in Cross-Plane Direction of Superlattices," Phys. Rev. B, **57**, pp. 14958–14973.
[72] Hyldgaard, P., and Mahan, G. D., 1997, "Phonon Superlattice Transport," Phys. Rev. B, **56**, pp. 10754–10757.
[73] Tamura, S., Tanaka, Y., and Maris, H. J., 1999, "Phonon Group Velocity and Thermal Conduction in Superlattices," Phys. Rev. B, **60**, pp. 2627–2630.
[74] Bies, W. E., Radtke, R. J., and Ehrenreich, H., 2000, "Phonon Dispersion Effects and the Thermal Conductivity Reduction in GaAs/AlAs Superlattices," J. Appl. Phys., **88**, pp. 1498–1503.
[75] Kiselev, A. A., Kim, K. W., and Stroscio, M. A., 2000, "Thermal Conductivity of Si/Ge Superlattices: A Realistic Model with a Diatomic Unit Cell," Phys. Rev. B, **62**, pp. 6896–6899.
[76] Yang, B., and Chen, G., 2001, "Anisotropy of Heat Conduction in Superlattices," Microscale Thermophys. Eng., **5**, pp. 107–116.
[77] Volz, S. G., Saulnier, J. B., Chen, G., and Beauchamp, P., 2000, "Computation of Thermal Conductivity of Si/Ge Superlattices by Molecular Dynamics Techniques," Microelectronics J., **31**, pp. 815–819.
[78] Liang, X. G., and Shi, B., 2000, "Two-Dimensional Molecular Dynamics

Simulation of the Thermal Conductance of Superlattices with Lennard-Jones Potential," Mater. Sci. Eng., **292**, pp. 198–202.
[79] Chen, G., 1999, "Phonon Wave Heat Conduction in Thin Films and Superlattices," ASME J. Heat Transfer , **121**, pp. 945–953.
[80] Balandin, A., and Wang, K. L., 1998, "Effect of Phonon Confinement on the Thermoelectric Figure of Merit of Quantum Wells," J. Appl. Phys., **84**, pp. 6149–6153.
[81] Ju, Y. S., and Goodson, K. E., 1999, "Phonon Scattering in Silicon Films with Thickness of Order 100 nm," Appl. Phys. Lett., **74**, pp. 3005–3007.
[82] Ziman, J. M., 1960, *Electrons and Phonons*, Clarendon Press, Oxford.
[83] Venkatasubramian, R., Siivola, E., and Colpitts, T. S., 1998, "In-Plane Thermoelectric Properties of Freestanding Si/Ge Superlattice Structures," *Proc. Int. Conf. Thermoelectrics*, ICT'98, pp. 191–197.
[84] Simkin, M. V., and Mahan, G. D., 2000, "Minimum Thermal Conductivity of Superlattices," Phys. Rev. Lett., **84**, pp. 927–930.
[85] Arutyunyan, L. I., Bogomolov, V. N., Kartenko, N. F., Kurdyukov, D. A., Popov, V. V., Prokof'ev, A. V., Smirnov, I. A., and Sharenkova, N. V., 1997, "Thermal Conductivity of a New Type of Regular-Structure Nanocomposites: PbSe in Opal Pores," Phys. Solid State, **39**, pp. 510–514.
[86] Song, D. W., Shen, W.-N., Zeng, T., Liu, W. L., Chen, G., Dunn, B., Moore, C. D., Goorsky, M. S., Radetic, R., and Gronsky, R., 1999, "Thermal Conductivity of Nano-Porous Bismuth Thin Films for Thermoelectric Applications," ASME HTD-Vol. 364-1, pp. 339–344.
[87] Volz, S. G., and Chen, G., 1999, "Molecular Dynamics Simulation of Thermal Conductivity of Si Nanowires," Appl. Phys. Lett., **75**, pp. 2056–2058.
[88] Walkauskas, S. G., Broido, D. A., Kempa, K., and Reinecke, T. L., 1999, "Lattice Thermal Conductivity of Wires," J. Appl. Phys., **85**, pp. 2579–2582.
[89] Kim, P., Shi, L., Majumadar, A., and McEuen, P. L., 2001, "Thermal Transport Measurements on Individual Multi-walled Nanotubes," Phys. Rev. Lett., **87**, p. 215502.
[90] Cahill, D. G., 1990, "Thermal Conductivity Measurement from 30-K to 750-K—The 3-Omega Method," Rev. Sci. Instrum., **61**, pp. 802–808.
[91] Lee, S. M., and Cahill, D. G., 1997, "Heat Transport in Thin Dielectric Films," J. Appl. Phys., **81**, pp. 2590–2595.
[92] Borca-Tasciuc, T., Kumar, R., and Chen, G., 2001, "Data Reduction in 3ω Method for Thin Film Thermal Conductivity Measurements," Rev. Sci. Instrum., **72**, pp. 2139–2147.
[93] Cahill, D. G., Fischer, H. E., Klitsner, T., Swartz, E. T., and Pohl, R. O., 1989, "Thermal Conductivity of Thin Films: Measurement and Understanding," J. Vac. Sci. Technol. A, **A7**, pp. 1259–1266.
[94] Hatta, I., 1990, "Thermal Diffusivity Measurement of Thin Films and Multilayered Composites," Int. J. Thermophys., **11**, pp. 293–303.
[95] Volklein, F., and Starz, T., 1997, "Thermal Conductivity of Thin Films—Experimental Methods and Theoretical Interpretation," *Proc. Int. Conf. Thermoelectrics*, ICT'97, pp. 711–718.
[96] Goodson, K. E., and Ju, Y. S., 1999, "Heat Conduction in Novel Electronic Films," Annu. Rev. Mater. Sci., **29**, pp. 261–293.
[97] Yang, B., Liu, J. L., Wang, K. L., and Chen, G., 2002, "Simultaneous Measurements of Seebeck Coefficient and Thermal Conductivity Across Superlattices," Appl. Phys. Lett., in press.
[98] Fleurial, J.-P., Snyder, G. J., Herman, J. A., Giauque, P. H., et al. 1999, "Thick-Film Thermoelectric Microdevices," *Proc. Int. Conf. Thermoelectrics*, ICT'99, pp. 294–300.
[99] Kishi, M., Yoshida, Y., Okano, H., Nemoto, H., Funanami, Y., Yamamoto, M., and Kanazawa, H., 1997, "Fabrication of a Miniature Thermoelectric Module with Elements Composed of Sintered Bi-Te Compounds," *Proc. Int. Conf. Thermoelectrics*, ICT'97, pp. 653–656.
[100] Rushing, L., Shakouri, A., Abraham, P., and Bowers, J. E., 1997, "Micro Thermoelectric Coolers for Integrated Applications," *Proc. Int. Conf. Thermoelectrics*, ICT'97, pp. 646–649.
[101] Volklein, F., Blumers, M., and Schmitt, L., "Thermoelectric Microsensors and Microactuators (MEMS) Fabricated by Thin Film Technology and Micromachining," *Proc. Int. Conf. Thermoelectrics*, ICT'99, pp. 285–293.
[102] Yao, D.-J., Kim, C.-J., and Chen, G., 2000, "Design of Thermoelectric Thin Film Coolers," ASME HTD-Vol. 366-2, pp. 245–251.
[103] Dutta, N. K., Cella, T., Brown, R. L., and Huo, D. T. C., 1985, "Monolithically Integrated Thermoelectric Controlled Laser Diode," Appl. Phys. Lett., **47**, pp. 222–224.
[104] Chen, G., 1996, "Heat Transfer in Micro- and Nanoscale Photonic Devices," Annu. Rev. Heat Transfer, **7**, pp. 1–57.
[105] Piprek, J., Akulova, Y. A., Babic, D. I., Coldren, L. A., and Bowers, J. E., 1998, "Minimum Temperature Sensitivity of 1.55 μm Vertical-Cavity Lasers at 30 nm Gain Offset," Appl. Phys. Lett., **72**, pp. 1814–1816.
[106] Berger, P. R., Dutta, N. K., Choquette, K. D., Hasnain, G., and Chand, N., 1991, "Monolithically Peltier-Cooled Vertical-Cavity Surface-Emitting Lasers," Appl. Phys. Lett., **59**, pp. 117–119.
[107] Corser, T. A., 1991, "Qualification and Reliability of Thermoelectric Coolers for Use in Laser Modules," 41st Electronic Components and Technology Conference, Atlanta, GA, USA, May, pp. 150–156.
[108] Cheng, Y.-K., Tsai, C.-H., Teng, C.-C., and Kang, S.-M., 2000, *Electrothermal Analysis of VLSI Systems*, Kluwer Academic Publishers.
[109] Kishi, M., Nemoto, H., Hamao, T., Yamamoto, M., Sudou, S., Mandai, M., and Yamamoto, S., 1999, "Micro Thermoelectric Modules and Their Application to Wristwatches as an Energy Source," *Proc. Int. Conf. Thermoelectrics*, ICT'99, pp. 301–307.

Thermal conductivity of Si/SiGe and SiGe/SiGe superlattices

Scott T. Huxtable, Alexis R. Abramson, Chang-Lin Tien, and Arun Majumdar[a)]
Department of Mechanical Engineering, University of California, Berkeley, California 94720

Chris LaBounty, Xiaofeng Fan, Gehong Zeng, and John E. Bowers
Department of Electrical and Computer Engineering, University of California, Santa Barbara, California 93106

Ali Shakouri
Jack Baskin School of Engineering, University of California, Santa Cruz, California 95064

Edward T. Croke
HRL Laboratories, LLC, Malibu, California 90265

(Received 13 November 2001; accepted for publication 20 December 2001)

The cross-plane thermal conductivity of four $Si/Si_{0.7}Ge_{0.3}$ superlattices and three $Si_{0.84}Ge_{0.16}/Si_{0.76}Ge_{0.24}$ superlattices, with periods ranging from 45 to 300 and from 100 to 200 Å, respectively, were measured over a temperature range of 50 to 320 K. For the $Si/Si_{0.7}Ge_{0.3}$ superlattices, the thermal conductivity was found to decrease with a decrease in period thickness and, at a period thickness of 45 Å, it approached the alloy limit. For the $Si_{0.84}Ge_{0.16}/Si_{0.76}Ge_{0.24}$ samples, no dependence on period thickness was found and all the data collapsed to the alloy value, indicating the dominance of alloy scattering. This difference in thermal conductivity behavior between the two superlattices was attributed to interfacial acoustic impedance mismatch, which is much larger for $Si/Si_{0.7}Ge_{0.3}$ than for $Si_{0.84}Ge_{0.16}/Si_{0.76}Ge_{0.24}$. The thermal conductivity increased slightly up to about 200 K, but was relatively independent of temperature from 200 to 320 K.
 [DOI: 10.1063/1.1455693]

The thermal conductivity of semiconductor superlattice films is of great importance for the performance of thermoelectric coolers and power converters as well as of optoelectronic devices. Low thermal conductivity materials are desired for thermoelectric elements in order to reduce the backflow of heat from the hot to cold junctions, while high conductivity materials are necessary for optoelectronic devices in order to dissipate heat. Recently, the thermal conductivity of several superlattice systems including GaAs/AlAs,[1–5] Si/Ge,[6,7] Si/SiGe,[8] and Bi_2Te_3/Sb_2Te_3 (Refs. 9 and 10) has been a topic of significant interest. Several of these studies have shown that the thermal conductivity of a superlattice can be lower than a corresponding value calculated from Fourier heat conduction theory using the constituent materials' bulk thermal conductivity. However, the causes for this dramatic reduction in thermal conductivity are not entirely clear, and a variety of mechanisms such as mismatch in acoustic impedance and phonon spectra, miniband formation, and interface scattering due to roughness and defects have been proposed as contributors. The primary goal of this work is to examine the role of the interface and acoustic impedance mismatch in the reduction of the thermal conductivity.

The superlattice samples were grown by molecular beam epitaxy (MBE) at HRL Laboratories, LLC, and complete details of the growth can be found elsewhere.[11] They were grown on Si substrates upon which a buffer layer was deposited to reduce strain and to ensure good growth conditions for the superlattice. Capping the superlattice were Si_xGe_{1-x} alloys doped appropriately to provide good ohmic contact. The thickness ratio of the two superlattice materials within each period was maintained at 2:1 and 1:1 for the $Si/Si_{0.7}Ge_{0.3}$ and $Si_{0.84}Ge_{0.16}/Si_{0.76}Ge_{0.24}$ superlattices, respectively. Four $Si/Si_{0.7}Ge_{0.3}$ superlattices with period thicknesses of 300, 150, 75, and 45 Å were investigated along with three $Si_{0.84}Ge_{0.16}/Si_{0.76}Ge_{0.24}$ superlattices with period thicknesses of 100, 150, and 200 Å. This comparison offers an opportunity to examine how the thermal conductivity may be influenced by alloy and defect scattering, miniband formation, tunneling, and phonon spectra acoustic impedance mismatch (AIM). The $AIM=(\rho V)_1/(\rho V)_2$ where ρ is the material density, V is the speed of sound, and subscripts 1 and 2 represent the adjacent layers. The AIM partially determines the fraction of phonons reflected at each interface.[12]

The thermal conductivity of the superlattices was measured using the 3-ω technique.[13] In short, the 3-ω method is an ac technique in which a thin metal line is patterned on the surface of the sample and is utilized as both a heater and a thermometer. Since the capping layers are electrically conducting, a thin insulating layer (~100 nm thick SiO_2) is deposited between the cap and metal heater line. Detailed explanations of the 3-ω technique[13,14] and the fabrication of the insulating layer and heater/thermometer lines on the samples[15] can be found elsewhere. The 3-ω measurement directly gives the total thermal resistance of the superlattice, buffer, cap, and insulating layer. The individual contributions of the buffer, cap, and insulating layers are measured separately and subtracted, leaving only the thermal conductivity of the superlattice. In order to confirm the suitability of assuming one-dimensional heat flow (cross-plane thermal conductivity), heater lines of varying widths (16, 20, and 25 μm)

a)Electronic mail: majumdar@me.berkeley.edu

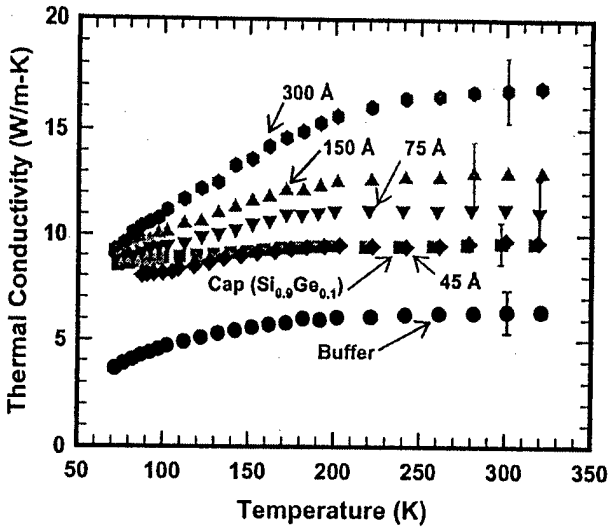


FIG. 1. Cross-plane thermal conductivity of four $Si/Si_{0.7}Ge_{0.3}$ superlattices. All superlattices are 3 μm thick and have a ratio of Si to $Si_{0.7}Ge_{0.3}$ of 2:1. The labels on the upper four curves refer to the period thickness. The curve labeled "Cap" is a 3.5 μm thick $Si_{0.9}Ge_{0.1}$ film, which has the same composition as the capping layers on the superlattices. The "Buffer" is a 1 μm thick $Si_{0.9}Ge_{0.1}$ film followed by a 1 μm thick $Si_{0.9}Ge_{0.1}/Si_{0.845}Ge_{0.15}C_{0.005}$ superlattice.

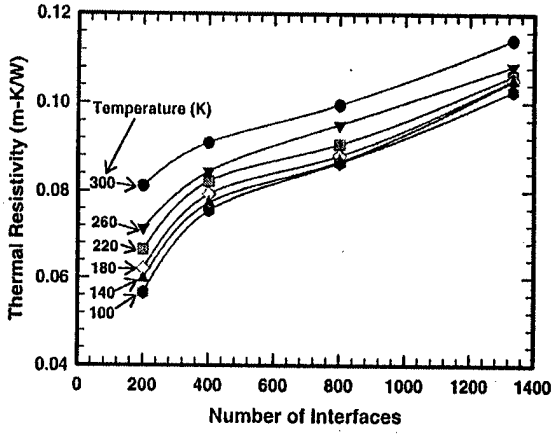


FIG. 2. Thermal resistivity of the four $Si/Si_{0.7}Ge_{0.3}$ superlattices as a function of the number of interfaces at six different temperatures (K). The resistivity decreases nearly linearly between 1334 and 400 interfaces, while it decreases slightly more between 400 and 200 interfaces.

were tested. The difference among thermal conductivity values measured with lines of these widths was small, indicating that the lines were adequately wide to assume one-dimensional heat flow.

To assess the reliability of our data, we first measured the thermal conductivity of a 3.5 μm thick $Si_{0.9}Ge_{0.1}$ film that was boron doped with a carrier concentration of 5×10^{19} cm^{-3} (see Fig. 1). This measurement was then compared with data recently reported by Yonenaga *et al.*[16] for a similarly composed $Si_{0.9}Ge_{0.1}$ sample, also boron doped, only with a slightly higher carrier concentration of 8×10^{19} cm^{-3}. Our measured value of 9.6 W/m K for room temperature thermal conductivity compared favorably with their value of 9.35 W/m K, thus confirming the reliability of our measurements.

The number of interfaces present in a $Si/Si_{0.7}Ge_{0.3}$ superlattice strongly influences the thermal conductivity as demonstrated in Fig. 1. This plot was generated using the data from the 20 μm wide heater lines, which should be closely representative of the cross-plane thermal conductivity. For a decrease in period thickness (an increase in the number of interfaces per unit length), there was a corresponding decrease in the thermal conductivity. For the 45 Å period thickness, the thermal conductivity approached that of the $Si_{0.9}Ge_{0.1}$ alloy. It is noteworthy that for a $Si/Si_{0.7}Ge_{0.3}$ superlattice with a 2:1 thickness ratio, the overall ratio of Si to Ge is 9:1, which is the same ratio within the alloy. It is also worth noting that the thermal conductivity never fell below the value of the corresponding $Si_{0.9}Ge_{0.1}$ alloy. This is in contrast to the observations of Lee *et al.*[6] on superlattices of Si/Ge. Presumably, the larger AIM of the Si/Ge superlattices along with the presence of misfit dislocations and interface defects in their case increased phonon scattering. In our particular case, however, the superlattice acts as much like an alloy of its constituents as a layered structure. All superlattices show similar temperature dependence: a gradual increase in thermal conductivity with temperature up to about 200 K, and only a slight increase from 200 to 320 K.

Figure 2 shows a plot of the thermal resistivity of the superlattices as a function of the number of interfaces. The resistivity decreases nearly linearly from 1334 to 400 interfaces, and then drops off somewhat more rapidly between 400 and 200 interfaces. The linear portion seemingly indicates that the additional interfaces simply add a corresponding thermal boundary resistance to the superlattices. However, the reason for the deviation from linear behavior for long periods is unclear and may result from the interplay between competing thermal mechanisms. The overall trend towards lower thermal conductivity for shorter periods eliminates the possibility of miniband formation being a dominant mechanism for heat transport in these superlattices for the following reasons. When miniband formation occurs in the phonon dispersion relation, the number of locations in wave-vector space where stop bands, or phonon band gaps, occur decreases with a decrease in period thickness, leading to fewer locations where there is a decrease in the phonon group velocity. Therefore, if miniband formation were domi-

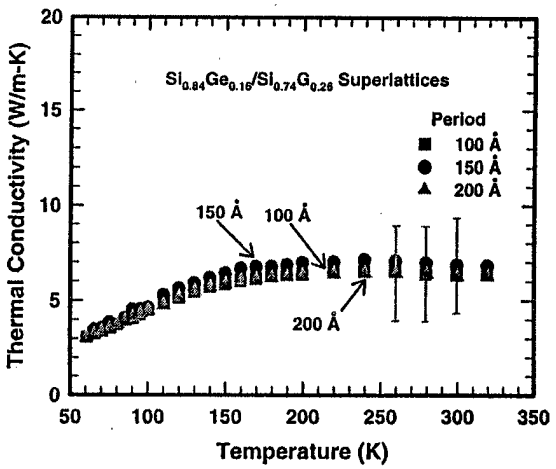


FIG. 3. Cross-plane thermal conductivity of three $Si_{0.84}Ge_{0.16}/Si_{0.76}Ge_{0.24}$ superlattices. All superlattices are 3 μm thick and the labels on the curves refer to the period thickness. The high degree of uncertainty is only in regard to the absolute value of the thermal conductivity; all three samples have essentially the same conductivity.

nant, decreasing the thickness of the periods would actually result in an increase in the thermal conductivity.[17]

The thermal conductivity of the $Si_{0.84}Ge_{0.16}/Si_{0.76}Ge_{0.24}$ superlattices measured with 25 μm heater lines is shown in Fig. 3. The main source of uncertainty in these data is due to the fact that a separate buffer sample was not available for measurement. Since the buffer layer consisted of a 0.95 μm thick layer of $Si_{0.8}Ge_{0.2}$ followed by a 3 μm thick graded SiGe film (graded from pure Si to $Si_{0.8}Ge_{0.2}$), we assumed that it acted thermally like a $Si_{0.9}Ge_{0.1}$ alloy that was measured previously. Nonetheless, the relative uncertainty among these three superlattices is not affected by this estimate. Unlike the Si/alloy samples, where the period thickness was significant with regard to the thermal conductivity, the alloy/alloy superlattice thermal conductivity was independent of the period thickness. We attribute this to the dissimilarity among the AIM of the superlattices. Using the data for the elastic constants of Si (Ref. 18) and Ge (Ref. 18) and a weighted average of data for the alloys,[19] and then averaging over all of the phonon modes and propagation directions, an estimate of an "average" AIM for each superlattice was obtained. The calculated average acoustic impedance mismatches were ~1.15 and ~1.03 for the $Si/Si_{0.7}Ge_{0.3}$ and $Si_{0.84}Ge_{0.16}/Si_{0.76}Ge_{0.24}$ superlattices, respectively. From Figs. 1–3, we saw that the number of interfaces had a significant effect on the heat transport in the Si/alloy superlattices, while the alloy/alloy superlattices showed no such effect. If the AIM were dominant, then one would expect that, for the alloy/alloy superlattice, the number of interfaces would have little, if any, influence on the thermal conductivity. Rather, alloy scattering appears to be the dominant mechanism. Furthermore, in comparing Figs. 1 and 3, we see that the alloy/alloy superlattice exhibits a slightly lower thermal conductivity than the Si/alloy superlattices.

In summary, the cross-plane thermal conductivities of $Si/Si_{0.7}Ge_{0.3}$ and $Si_{0.84}Ge_{0.16}/Si_{0.76}Ge_{0.24}$ superlattices were examined and compared. For the Si/alloy superlattice, the role of the interface in affecting the thermal conductivity is important; however, for the alloy/alloy superlattice it is not. Taking advantage of both alloy scattering and the influence of the AIM, future investigation of a high AIM alloy/alloy superlattice could demonstrate the combined effect on the thermal behavior of these structures.

[1] T. Yao, Appl. Phys. Lett. **51**, 1798 (1987).
[2] W. S. Capinski and H. J. Maris, Physica B **219 & 220**, 699 (1996).
[3] W. S. Capinski, H. J. Maris, T. Ruf, M. Cardona, K. Ploog, and D. S. Katzer, Phys. Rev. B **59**, 8105 (1999).
[4] G. Chen, C. L. Tien, X. Wu, and J. S. Smith, J. Heat Transfer **116**, 325 (1994).
[5] X. Y. Yu, G. Chen, A. Verma, and J. S. Smith, Appl. Phys. Lett. **67**, 3554 (1995).
[6] S.-M. Lee, D. G. Cahill, and R. Venkatasubramanian, Appl. Phys. Lett. **70**, 2957 (1997).
[7] T. Borca-Tasciuc, W. Liu, J. Liu, T. Zeng, D. W. Song, C. D. Moore, G. Chen, K. L. Wang, and M. S. Goorsky, Superlattices Microstruct. **28**, 199 (2000).
[8] G. Chen, S. Q. Zhou, D.-Y. Yao, C. J. Kim, X. Y. Zheng, Z. L. Liu, and K. L. Wang, 17th International Conference on Thermoelectrics, 1998, p. 202.
[9] R. Venkatasubramanian, Phys. Rev. B **61**, 3091 (2000).
[10] M. N. Touzelbaev, P. Zhou, R. Venkatasubramanian, and K. E. Goodson, J. Appl. Phys. **90**, 763 (2001).
[11] X. Fan, G. Zeng, E. Croke, G. Robinson, C. LaBounty, A. Shakouri, and J. E. Bowers, Phys. Low-Dimens. Semicond. Struct. **5/6**, 1 (2000).
[12] E. T. Swartz and R. O. Pohl, Rev. Mod. Phys. **61**, 605 (1989).
[13] D. G. Cahill, M. Katiyar, and J. R. Abelson, Phys. Rev. B **50**, 6077 (1994).
[14] D. G. Cahill, Rev. Sci. Instrum. **61**, 802 (1990).
[15] S. T. Huxtable, A. R. Abramson, A. Majumdar, C. L. Tien, C. LaBounty, X. Fan, G. Zeng, J. E. Bowers, A. Shakouri, and E. T. Croke, Proceedings of the ASME IMECE Conference, New York (in press).
[16] I. Yonenaga, T. Akashi, and T. Goto, J. Phys. Chem. Solids **62**, 1313 (2001).
[17] M. V. Simkin and G. D. Mahan, Phys. Rev. Lett. **84**, 927 (2000).
[18] A.-B. Chen and A. Sher, *Simiconductor Alloys: Physics and Materials Engineering* (Plenum, New York, 1995).
[19] V. T. Bublik, S. S. Gorelik, A. A. Zaitsev, and A. Y. Polyakov, Phys. Status Solidi B **66**, 427 (1974).

Electron transport through strongly coupled AlInP/GaInP superlattices

R. E. Martínez II[a)]
Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138

I. Appelbaum
Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138 and Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

C. V. Reddy, R. Sheth, and K. J. Russell
Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138

V. Narayanamurti
Division of Engineering and Applied Sciences and Department of Physics, Harvard University, Cambridge, Massachusetts 02138

J.-H. Ryou, U. Chowdhury, and R. D. Dupuis
Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, Texas 78712

(Received 5 August 2002; accepted 13 September 2002)

Using ballistic-electron-emission spectroscopy, electron transport through the principal (Γ_c,L_c) miniband of an $(Al_{0.5}In_{0.5}P)_{11}/(Ga_{0.5}In_{0.5}P)_{10}$ superlattice in the strong-coupling regime has been observed. Second derivative spectra of experimental data and Monte Carlo simulations were in agreement. © *2002 American Institute of Physics.* [DOI: 10.1063/1.1519350]

For optoelectronic applications, a highly efficient, stoichiometrically modifiable alternative to the well-established ternary arsenides, $Al_xGa_{1-x}As$, is desirable. The binary to quaternary phosphides, $(Al_xGa_{1-x})_yIn_{1-y}P$, have the second widest direct band gap of III–V semiconductors—exceeded only by the ternary nitrides, $Ga_xIn_{1-x}N$— with the advantage of GaAs lattice matching and luminescence in the visible range.[1,2] The ternary to quaternary $(Al_xGa_{1-x})_yIn_{1-y}P$/GaAs heterostructures have proven to be effective alternatives to $Al_xGa_{1-x}As$/GaAs heterostructures as field-effect transistors,[3] heterojunction bipolar transistors,[4] highly efficient (≤30%) tandem solar cells,[5] and vertical-cavity surface-emitting diodes.[6] The superiority of $(Al_xGa_{1-x})_yIn_{1-y}P$ over the $Al_xGa_{1-x}As$ analogues is due to the excellent characteristics found in phosphide systems—namely, large valence ($Ga_xIn_{1-x}P$) and conduction ($Al_xIn_{1-x}P$) band offsets, greater output power, higher operating temperatures, and lower oxidation and temperature degradation rates.[2,7]

Although many redeeming characteristics of the phosphide systems are limited in the arsenide systems, an unfortunate fact remains of the active region regardless of composition: large carrier leakage at room-temperature operation. One method of reducing this leakage is by increasing carrier recombination by enhancing confinement within the active region. Standard superlattices and strain-modulated aperiodic superlattice heterobarriers have been used to reduce leakage, as in thermalization losses, and enhance carrier confinement in light-emitting devices.[8] Incorporating the ternary phosphides, specifically $Al_xIn_{1-x}P$ and $Ga_xIn_{1-x}P$, into superlattices has been achieved with success;[9] however, transport through the phosphide superlattices is not as well understood as that of the arsenide superlattices.[10]

With a deeper understanding of transport in $(Al_xIn_{1-x}P)_n/(Ga_xIn_{1-x}P)_m$ superlattices, further promising facets of the phosphide-related, optoelectronic avenue will undoubtedly unmask. In this letter, we report the observation of electron transport through the principal (Γ_c,L_c) miniband of the strongly coupled $(Al_{0.5}In_{0.5}P)_{11}/(Ga_{0.5}In_{0.5}P)_{10}$ superlattice using ballistic-electron-emission spectroscopy.[11] This claim is supported by Monte Carlo simulations, which are in agreement with experimental data.

Ballistic-electron-emission spectroscopy (BEES) is a scanning tunneling microscopy-related method that allows for field-free carrier injection, a necessity when concerned with field-sensitive phenomena, such as resonant tunneling through quasibound states and miniband states. From the tip, forward-directed, hot electrons tunnel through the vacuum barrier and travel ballistically through the thin base metal. At the nonepitaxial metal–semiconductor interface, electrons scatter into available states of the semiconductor conduction band or reflect into the base metal. Nonepitaxial interfaces, phonon scattering, and, if present, buried heterostructure transmission probability, contribute to attenuation reducing the collector current ten to 100-fold. The base current is held constant, while the bias voltage is increased within a given range. The collector current is then recovered, typically, on the order of 10–100 pA and significantly larger than noise (1 pA). A detailed description and recent review of BEES has been reported elsewhere.[12]

The $(Al_{0.5}In_{0.5}P)_{11}/(Ga_{0.5}In_{0.5}P)_{10}$ superlattices were grown using low-pressure, metal–organic vapor deposition on n-GaAs:Si ⟨100⟩ substrates. Adduct-purified triethylgallium ($Ga[C_2H_5]_3$), trimethylindium ($In[CH_3]_3$), and trimethylaluminum ($Al[CH_3]_3$) were used as group III elements, and phosphine (PH_3) and arsine (AsH_3) were used as group

a)Author to whom correspondence should be addressed; electronic mail: martinez@deas.harvard.edu

TABLE I. Semiconductor parameters of GaAs, $Al_{0.5}In_{0.5}P$, and $Ga_{0.5}In_{0.5}P$ used in the simulations are tabulated. These include Γ_c and L_c energy differences, conduction band offsets, and effective masses self-consistently calculated at 298 K.

Parameter (298 K)	Suggested value
$\Delta\mathcal{E}^{\mathcal{E}_F-\Gamma}_{Au-GaAs}$	0.900 eV
$\Delta\mathcal{E}^{\Gamma-L}_{GaAs}$	0.285 eV
$\Delta\mathcal{E}^{\Gamma-L}_{Al_{0.5}In_{0.5}P}$	0.161 eV
$\Delta\mathcal{E}^{\Gamma-L}_{Ga_{0.5}In_{0.5}P}$	0.326 eV
$\Delta\mathcal{E}^{\Gamma-\Gamma}_{GaAs-Al_{0.5}In_{0.5}P}$	0.624 eV
$\Delta\mathcal{E}^{L-L}_{GaAs-Al_{0.5}In_{0.5}P}$	0.501 eV
$\Delta\mathcal{E}^{\Gamma-\Gamma}_{GaAs-Ga_{0.5}In_{0.5}P}$	0.186 eV
$\Delta\mathcal{E}^{L-L}_{GaAs-Ga_{0.5}In_{0.5}P}$	0.227 eV
m^{Γ}_{GaAs}	0.067 m_e
m^{L}_{GaAs}	0.560 m_e
$m^{\Gamma}_{Al_{0.5}In_{0.5}P}$	0.110 m_e
$m^{L}_{Al_{0.5}In_{0.5}P}$	0.139 m_e
$m^{\Gamma}_{Ga_{0.5}In_{0.5}P}$	0.092 m_e
$m^{L}_{Ga_{0.5}In_{0.5}P}$	0.114 m_e

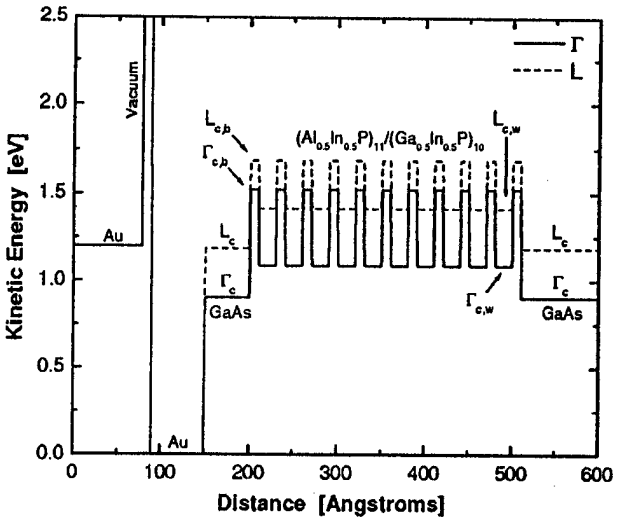


FIG. 1. The BEES transport diagram from the tip through the conduction band diagram of an $(Al_{0.5}In_{0.5}P)_{11}/(Ga_{0.5}In_{0.5}P)_{10}$ superlattice is shown with an incident electron kinetic energy of 1.20 eV, corresponding to a tip bias of −1.20 V. Valley minima of GaAs, $Al_{0.5}In_{0.5}P$, and $Ga_{0.5}In_{0.5}P$ are labeled by Γ_c and L_c, $\Gamma_{c,b}$ and $L_{c,b}$ (barrier), and $\Gamma_{c,w}$ and $L_{c,w}$ (well), respectively.

V elements. The epitaxial growth was initiated with 500 Å of an undoped GaAs buffer layer, followed by a cladding layer of 1 nm $Al_{0.5}In_{0.5}P$. A set of 20 layers was grown in the order of 2 nm of $Ga_{0.5}In_{0.5}P$ (10), then 1 nm of $Al_{0.5}In_{0.5}P$ (10), followed by cap layer of 50 Å of undoped GaAs. Ohmic (collector) contacts of a 1 μm Au/Ge/Ni alloy were deposited by electron beam evaporation, followed by annealing in high vacuum ($\sim 10^{-7}$ Torr) at 550 °C for 300 s. The Schottky (base) contacts were prepared by thermal deposition of 60 Å Au in high vacuum ($\sim 10^{-7}$ Torr). BEES current–voltage spectra were acquired under nitrogen flow at room temperature, while the tunneling current was held constant at approximately 4 nA over the tip bias range of −0.8 to −1.8 V. To ensure uniformity, approximately 200 000 scans were taken from different points over the base and averaged.

Semiempirical calculations of electron transport were performed using a semiclassical, momentum-space Monte Carlo (MC) simulation. Transmission coefficients from tip to base metal through vacuum ($\mathcal{T}_{t\text{-}m}$), base metal ($\mathcal{T}_m$), nonepitaxial (base) metal–semiconductor interface ($\mathcal{T}_{m\text{-}s}$), cap layer, heterostructure, and buffer layer interfaces ($\mathcal{T}_s$) were included in the simulations, in addition to thermal effects ($\mathcal{T}_{\text{temp}}$) and elastic/inelastic s-wave scattering. Varshni and Kane (eight-band $\mathbf{k}\cdot\mathbf{p}$ perturbation theory) parametrizations were used to model temperature-dependent spherical, nonparabolic valleys and energy/temperature-dependent effective masses, respectively.[2] Previously published[12] and self-consistently calculated[2] values for parameters used are given in Table I. Intervalley tunneling, LO phonon coupling, and impact ionization were not included in the simulations.

A synopsis of simulated transport is as follows: In the tip, forward-directed electrons are randomly chosen in momentum space using the von Neumann acceptance/rejection method. Assuming parallel momentum and energy conservation during transport from the tip to the base, electron transmission through vacuum is modeled by the WKB approximation of a one-dimensional trapezoidal barrier potential. The base current is a product of this transmission probability and incident electron flux and kept approximately constant to reflect experimental conditions. Attenuation from inelastic electron–electron scattering in the base metal is included within the random phase approximation mean-free path.[12] Elastic s-wave scattering at the metal–semiconductor interface is modeled by a random reorientation of momentum and conservation of energy for a given scattering probability (0–1). The remaining forward-directed electrons with sufficient kinetic energy in the base metal scatter across the step potential and couple into available states with compatible parallel momenta in the first Brillouin zone of the semiconductor. To model the transmission from the cap layer through a heterostructure to the buffer layer, a transfer matrix calculation is applied for each conduction band. The collector current is then given by a summation over N states, including the product of each transmission/attenuation coefficient,

$$I_c \propto \sum_{j=1}^{N} k_j^{\perp t} \mathcal{T}_{\text{temp}}(\mathcal{E}_j^t - \mathcal{E}_F^t, \mathcal{E}_j^m - \mathcal{E}_F^m, T)\mathcal{T}_{t\text{-}m}(k_j^{\perp t})\mathcal{T}_m(\mathcal{E}_j^m - \mathcal{E}_F^m, T)\mathcal{T}_{m\text{-}s}(k_j^{\perp m}, k_j^{\perp s}, T)\mathcal{T}_s(\mathcal{E}_j^s, T),$$

where $k_j^{\perp t/m/s}$, $\mathcal{E}_j^{t/m/s}$, $\mathcal{E}_F^{t/m}$, and T are the perpendicular momentum and energy of the jth electron in the tip, base, or semiconductor, Fermi energy of the tip or base metal, and temperature, respectively. Finally, the normalized total current is calculated from the ratio of the collector and the base currents.

The conduction band diagram of an $(Al_{0.5}In_{0.5}P)_{11}/(Ga_{0.5}In_{0.5}P)_{10}$ superlattice is shown in Fig. 1. The second derivative[13] of the averaged BEES spectra of a reference epitaxial GaAs on an n-GaAs:Si ⟨100⟩ sample is shown in Fig. 2(a). Comparatively, the second derivative of the MC spectra of the reference epitaxial GaAs with a scattering probability of 0.90 is shown in Fig. 2(c).[12,14] At small tip bias, the simulation reproduces the experimental data with the correct energetic onset and features of Γ_c and L_c valley contributions; the slight discrepancy at large tip bias is

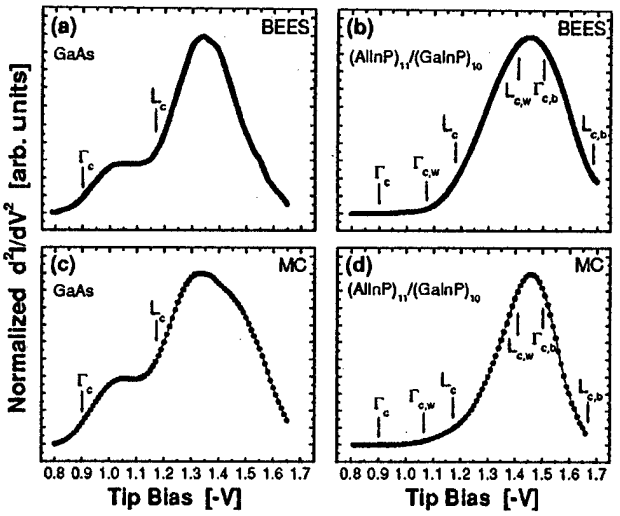


FIG. 2. (a)–(d) Numerical second derivatives of the BEES spectra and MC simulations of the reference epitaxial GaAs and the $(Al_{0.5}In_{0.5}P)_{11}/(Ga_{0.5}In_{0.5}P)_{10}$ superlattice are shown in (a) and (b) and (c) and (d), respectively. Valley minima are labeled for reference with Fig. 1.

consistent with the level of approximations considered in the MC simulations. With the incorporation of the superlattice, the transport threshold is shifted by 180 meV, as seen in Fig. 2(b). The second derivative of the MC spectra of the $(Al_{0.5}In_{0.5}P)_{11}/(Ga_{0.5}In_{0.5}P)_{10}$ superlattice shows a wide peak (full width at half maximum ~500 meV), beginning slightly above the $Ga_{0.5}In_{0.5}P$ Γ_c point minimum (1.10 eV) and terminating above the $Al_{0.5}In_{0.5}P$ L_c point minimum (1.70 eV), as seen in Fig. 2(d).[12] Due to strong nonepitaxial metal–semiconductor interface scattering,[14] the GaAs cap layer becomes an electron filter; that is, allowing mostly L_c and some Γ_c electrons to proceed, while effectively preventing the X_c electrons (with short mean-free path) from coupling with the adjacent X_c valley states of $Al_{0.5}In_{0.5}P$. Therefore, for consistency with experiment, the X_c valley is not included in the simulations of the reference epitaxial GaAs and $(Al_{0.5}In_{0.5}P)_{11}/(Ga_{0.5}In_{0.5}P)_{10}$ superlattice. The agreement between experiment and simulation confirm that the transport is dominated by the Γ_c and L_c valleys.

Electron transport in the principal (Γ_c,L_c) miniband of the strongly coupled $(Al_{0.5}In_{0.5}P)_{11}/(Ga_{0.5}In_{0.5}P)_{10}$ superlattice has been observed and discussed. Using BEES, transport through the superlattice was resolved at energies well below the $Al_{0.5}In_{0.5}P$ L_c barrier height—a purely quantum mechanical phenomenon. In support of this claim, MC simulations of electron transport were in agreement with the observed data.

The authors would like to thank K. B. Shelby, I. Shalish, N. Master, and I. B. Altfeder for helpful comments and insightful discussions. This research was funded in part by the National Science Foundation for the Materials Research Science and Engineering Center (MRSEC), National Science Foundation Grant Nos. ECS-9996093 and ECS-9906047 at Harvard University and Grant No. DMR-9612283 at the University of Texas at Austin, and a subgrant from the University of Texas at Austin to Harvard University.

[1] X. H. Zhang, S. J. Chua, and W. J. Fan, Appl. Phys. Lett. **73**, 1098 (1998).
[2] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, J. Appl. Phys. **89**, 5815 (2001).
[3] T. Ohori, M. Takechi, M. Suzuki, M. Takikawa, and J. Komeno, J. Cryst. Growth **93**, 905 (1988); T. Nittono, S. Sugitani, and F. Hyuga, J. Appl. Phys. **78**, 5387 (1995).
[4] T. Kobayashi, K. Taira, F. Nakamura, and H. Kawai, J. Appl. Phys. **65**, 4898 (1989); T. Nittono, S. Sugitani, and F. Hyuga, *ibid.* **78**, 5387 (1995).
[5] T. Takamoto, E. Ikeda, H. Kurita, and M. Ohmori, Appl. Phys. Lett. **70**, 381 (1996).
[6] M. J. Ries, N. Holonyak, Jr., E. I. Chen, S. A. Maranowski, M. R. Islam, A. L. Holmes, and R. D. Dupuis, Appl. Phys. Lett. **67**, 1107 (1995).
[7] M. Mannoh, J. Hoshina, S. Kamiyama, H. Ohta, Y. Ban, and K. Ohnaka, Appl. Phys. Lett. **62**, 1173 (1992); O. Shuler, O. Dehaese, X. Wallart, and F. Mollot, J. Appl. Phys. **84**, 765 (1998).
[8] M. R. Islam, R. V. Chelakara, and R. D. Dupuis, Appl. Phys. Lett. **67**, 2057 (1995).
[9] Y. Kawamura and H. Asahi, Appl. Phys. Lett. **45**, 152 (1984); M. J. Hafich, H. Y. Lee, G. Y. Robinson, D. Li, and N. Otsuka, J. Appl. Phys. **69**, 752 (1990).
[10] C. Rauch, G. Strasser, K. Unterrainer, E. Gornik, and B. Brill, Appl. Phys. Lett. **70**, 649 (1997); J. Smoliner, R. Heer, C. Eder, and G. Strasser, Phys. Rev. B **58**, R7516 (1998); R. Heer, J. Smoliner, G. Strasser, and E. Gornick, Appl. Phys. Lett. **73**, 3138 (1998).
[11] W. J. Kaiser and L. D. Bell, Phys. Rev. Lett. **60**, 1406 (1988).
[12] V. Narayanamurti and M. Kozhevnikov, Phys. Rep. **349**, 447 (2001).
[13] D. L. Smith and S. M. Kogan, Phys. Rev. B **54**, 10354 (1996).
[14] D. L. Smith, E. Y. Lee, and V. Narayanamurti, Phys. Rev. Lett. **80**, 2433 (1998); D. L. Smith, M. Kozhevnikov, E. Y. Lee, and V. Narayanamurti, Phys. Rev. B **61**, 13914 (2000).

Confinement-Enhanced Electron Transport across a Metal-Semiconductor Interface

I. B. Altfeder, J. A. Golovchenko, and V. Narayanamurti
Division of Engineering and Applied Science, Harvard University, Cambridge, Massachusetts 02138
and Department of Physics, Harvard University, Cambridge, Massachusetts 02138
(Received 26 December 2000; published 11 July 2001)

We present a combined scanning tunneling microscopy and ballistic electron emission microscopy study of electron transport across an epitaxial Pb/Si(111) interface. Experiments with a self-assembled Pb nanoscale wedge reveal the phenomenon of confinement-enhanced interfacial transport: a proportional increase of the electron injection rate into the semiconductor with the frequency of electron oscillations in the Pb quantum well.

DOI: 10.1103/PhysRevLett.87.056801 PACS numbers: 73.21.-b, 68.37.Ef, 73.40.-c, 81.07.-b

The physical properties of self-organized epitaxial films are important for the study of low-dimensional electron transport [1,2], electron confinement in self-assembled nanostructures [3–5], and geometry of the buried interfaces [6,7]. These films often consist of an extremely thin epitaxial wetting layer uniformly covering a semiconductor surface, and three-dimensional self-assembled nanostructures that grow on the top of this layer by a mechanism known as the Stranski-Krastanov growth mode. Among the remarkable features of these films is a continuous metallic surface conductivity, which allows one to electrically contact *individual* nanostructures to study their transport properties. Understanding the influence of electron confinement, i.e., resonant electron states and size-dependent electron energy quantization, on the transport properties in these nanostructures is of fundamental importance in the development of nanoscale electronic devices. The experimental study of such quantum transport phenomena in nanostructures can be done with a combination of scanning tunneling microscopy (STM) [8] and ballistic electron emission microscopy (BEEM) [9–12].

Being a three-terminal modification of STM, BEEM is a powerful method for the study of transport phenomena. In the BEEM experiment, the semiconductor surface is usually covered with a thin (~50 Å) metal layer, where the electrical potential is typically kept at zero. At low tip-sample bias, electrons tunneling from the STM tip spread through the metallic surface of the sample. However, when the energy of the tunneling electrons exceeds the Schottky barrier height, which is typically about 1 eV, some of these electrons are injected into the bulk of the semiconductor. The injected current is then collected by an Ohmic contact at the back side of the semiconductor substrate. BEEM methodology has been previously used in the study of Schottky barriers [9,13], semiconductor density of states [14], resonant transport through semiconductor heterostructures [14] and buried quantum dots [15], and defects at buried interfaces [16].

In this Letter, we report the first combined STM/BEEM experiment on extremely thin metal islands that self-assemble during epitaxial growth. The results show a proportional relationship between the ballistic transport across the epitaxial interface and the electron confinement energy in a metallic quantum well (QW), suggesting a new model for interfacial transport in this regime.

It has been previously shown that, on the Si(111) substrate, Pb islands begin to grow after an initial continuous double layer of Pb wets the surface [17]. On stepped substrates, Pb islands often possess a wedge-shaped geometry with an atomically flat top [4]. In this geometry, the metal thickness varies in a discrete manner due to atomic steps at the substrate-metal interface. Using these wedge-shaped nanostructures in conjunction with STM allows a convenient *in situ* adjustment of the metallic QW thickness. These properties make the Pb/Si nanostructures a perfect object for the study of both electron confinement effects and quantum transport phenomena.

The semiconductor substrate for our experiment was cut from a lightly doped *n*-type Si(111) wafer. An Ohmic back contact was fabricated *ex situ* by As implantation, which created a heavily doped *n*-type layer. The sample was placed into an ultrahigh vacuum (UHV) chamber with a base pressure of 1.5×10^{-10} Torr, where its surface was cleaned by a sequence of direct current heating (1100 °C) and ion-beam sputtering (Ne^+). The quality and the chemical composition of the surface were monitored by Auger electron spectroscopy and low-energy electron diffraction. Following the cleaning, Pb atoms were deposited from an effusion cell onto a 7×7 reconstructed Si(111) substrate. The sample was then *in situ* transferred to a UHV STM operating at room temperature.

The schematic of the experiment is presented in Fig. 1. The electrical contact to the surface Pb overlayer was made *in situ* using a spring contact located 1 mm apart from the STM tip. This contact was grounded to provide a separate conductance path for low-energy electrons. The current injected into the substrate, I_c, was measured with a picoammeter connected to the back side Ohmic contact held at zero potential. During the collector current vs tunnel bias (I_c-V) measurements, the tunneling current was maintained constant. The inset of Fig. 1 shows a typical I_c-V characteristic obtained on the Pb wetting layer. We found that, as long as the energy of tunneling electrons,

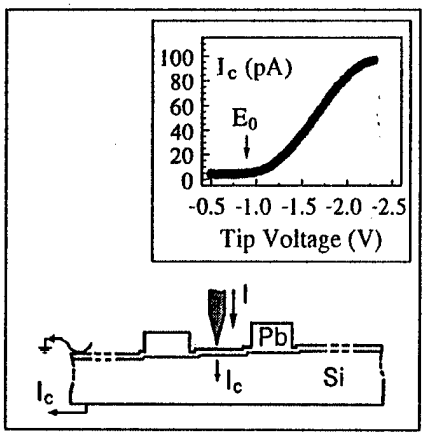


FIG. 1. Schematics of the *in situ* STM/BEEM experimental setup. Inset: Typical BEEM current-voltage characteristic obtained on the Pb wetting layer. The threshold $E_0 = 0.85$ eV for the electron injection into the bulk corresponds to the Pb/Si(111) Schottky barrier height.

$E = eV$, does not exceed the Pb/Si(111) Schottky barrier height [18] ($E_0 = 0.85 \pm 0.5$ eV), variations of I_c do not exceed the noise level. Obviously, at these energies all the tunneling current is collected from the 2D Pb wetting layer. Increasing the energy of the tunneling electrons above E_0 results in an increase of I_c, which is initially nearly quadratic and later evolves into a linear dependence on E [19]. This presents a clear evidence that the Pb wetting layer possesses a metallic conductivity, and, hence, can be used to electrically contact the nanostructures.

The complicated structure of our samples, consisting of objects of different dimensionality, raises a question as to how local the BEEM current is. In order to answer this question, we show in Fig. 2 simultaneously acquired STM and BEEM images (respectively) of our sample. In the STM image, we observe a significant density of surface steps. These steps, however, do not prevent the continuous

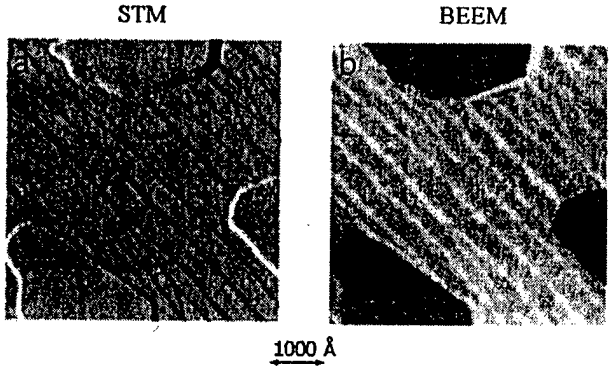


FIG. 2. 5000 × 5000 Å STM and BEEM images of epitaxial Pb(111) film on Si(111) acquired simultaneously at a tip voltage of −1.5 V. A significantly smaller BEEM current is observed within the borders of the Pb islands as compared to the surrounding Pb wetting layer. The gray scale in the BEEM image spans the range between 0 to 12 nA.

metallic conductivity of the Pb wetting layer, as the BEEM threshold E_0 is clearly observed throughout this layer. On top of the wetting layer, we observe 3D Pb islands with typical lateral sizes of 2000 Å and thickness varying from 30 to 80 Å. Analysis of the STM and BEEM images shows a significant reduction of the electron injection across the Pb islands as compared to the surrounding Pb wetting layer [20]. Apparently, most of the electrons, which tunnel into the Pb islands, thermalize below the Schottky barrier by the time they laterally diffuse into the surrounding wetting layer. As we show below, the electron injection into the semiconductor is laterally localized not only within a single Pb island but also within a single constant-height region of the wedge-shaped Pb islands.

In Fig. 3, we show a series of I_c-V characteristics obtained from Pb islands of different thickness. Since the interfacial structure remains intact below Pb islands [7], the Schottky threshold on the islands is the same as on the wetting layer. However, the shapes of these curves are different. An additional steplike feature appears in all the curves of Fig. 3, at an energy position that is clearly thickness dependent. The position of this steplike feature increases in energy as the film thickness decreases. This clearly suggests that the I_c-V features originate from the

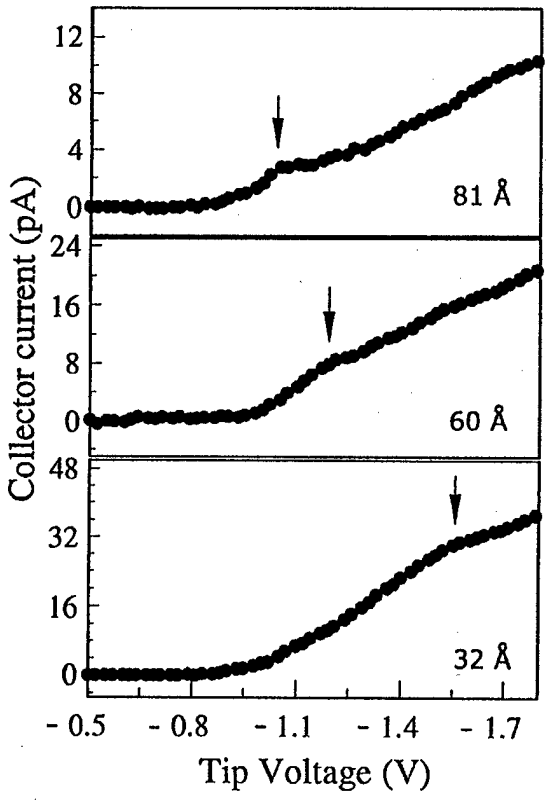


FIG. 3. BEEM current-voltage characteristics obtained on Pb islands of different thickness. The threshold, E_0, is the same as on the wetting layer. The additional thickness-dependent features that appear on these curves indicate resonant electron states in Pb islands.

resonant electron states in the Pb islands. A previous study of resonant states in Pb [4] with scanning tunneling spectroscopy has shown that the electron confinement energy,

$$\Delta = \frac{\pi \hbar v_F}{H}, \quad (1)$$

where v_F is the Fermi velocity in Pb, indeed increases as the metal thickness, H, decreases. However, unlike in tunneling spectra, where resonant steps repeat with an interval of Δ, the transmission (I_c-V) spectra always contain only one thickness-dependent steplike feature. As we shall show later, this is an inherent property of a three-terminal experiment.

In Fig. 4, we show the typical dependence of the electron injection rate (I_c/I) into the bulk Si(111) states on the thickness of Pb film. The data points in Fig. 4 were obtained by translating the STM tip over the top surface of the Pb *wedge* at a fixed tunneling current (10 nA) and a fixed voltage (−1.9 V) and measuring I_c for different metal thickness. We observe a continuous decrease of I_c with an increase of the metal thickness. Variations of I_c with the metal thickness have been attributed to electron attenuation over the mean-free path, λ, which implies the exponential dependence [21]: $I_c \propto \exp(-H/\lambda)$. This model, being valid only in the strong scattering limit, does not account for the *alternative* regime, i.e., development of standing electron waves in a metallic film [4,22]. Standing electron waves in the epitaxial Pb/Si(111) films have been directly observed in previous studies of this system [4,7]. In those studies, it was shown that the tunneling density of states in Pb islands exhibits oscillating energy dependence both below and above the Pb/Si Schottky barrier. The same experiment conducted at room temperature indicates a ballistic mean-free path of more than 100 Å.

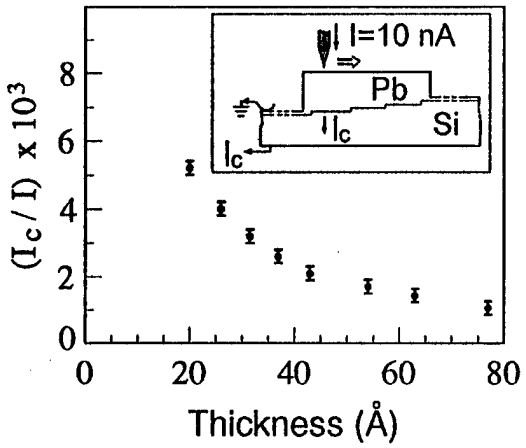


FIG. 4. Dependence of the normalized BEEM current, I_c/I, on the Pb thickness. The data points were obtained by translating the STM tip over the top surface on the self-assembled Pb wedge at a fixed tip voltage of −1.9 V and a fixed tunneling current of 10 nA (as shown in the inset). The errors are indicated by the vertical bars.

Therefore, the picture of a "particle in a box" seems to be appropriate for the explanation of the thickness-dependent ballistic transport across the Pb islands.

Consider an electron, of energy $E > E_0$, tunneling from a STM tip into a metal QW. The transversal motion of this electron is described by periodic oscillations across the QW with a frequency $\Omega = v/2H$, where $v \approx v_F$. Hence, the average charge injection into the semiconductor is given by

$$Q_c = \beta e(\Omega\tau), \quad (2)$$

where β is the interfacial transmission probability, e is the electron charge, and τ is the bulk electron relaxation time. The product βe represents the charge injection per attempt, while $\Omega\tau$ is the integral number of attempts to traverse the metal-semiconductor interface, which is given by the number of electron oscillations inside a QW that contribute to the interfacial transport. The rest of the charge can spread only through the wetting layer. As shown in a previous study, the low-energy part ($E < E_0$) of the tunneling current can be neglected at large negative tip voltage [4]. Thus, the ratio of the BEEM current to the tunneling current is given by the ratio between Q_c and e, and equals

$$\frac{I_c}{I} = \beta(\Omega\tau) \approx \frac{\beta v_F \tau}{2H}. \quad (3)$$

The inverse proportion of the BEEM current to the QW height, predicted in Eq. (3), can thus explain the decrease of I_c with the metal thickness. To show this dependence, we plot $(I_c/I)^{-1}$ as a function of H (Fig. 5). The resulting linear dependence is clearly observed in Fig. 5. Hence, comparing Eqs. (1) and (3), we infer that $I_c \propto \Delta$, i.e., the interfacial transport in epitaxial QWs is determined by the electron confinement energy.

Deviations from exponential dependence in thin Au film on Si(111) were reported earlier [21,23]. However, as bulk resonant states in Au films have never been observed, the

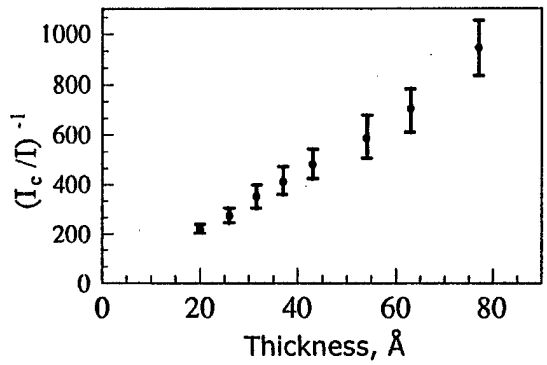


FIG. 5. The inverse of the normalized BEEM current, $(I_c/I)^{-1}$, plotted as a function of the Pb thickness. A proportional dependence is clearly observed. The errors are indicated by the vertical bars.

scaling predicted by Eq. (3) was not observed as well. Besides, as Au films do not grow in Stranski-Krastanov geometry, direct thickness measurement with STM is not possible. By contrast, the two noticeable features of the Pb/Si(111) interface are (i) self-assembly of the wedge-shaped islands, allowing for direct thickness determination, and (ii) electron focusing in the [111] direction due to the Fermi surface anisotropy, enhancing resonant electron states in the Pb QWs [24]. The combination of these two features allows one to consider a Pb island a model *quantum box* system.

As was first recognized by Bardeen [25], in two-terminal transport experiments, such as tunneling into a quantum box, the scaling effect cannot be manifested. Transport in these systems can be described as $M^2\rho$, where M is the tunneling matrix element, and ρ is the density of states. Because of $M \propto \psi \propto 1/\sqrt{H}$, where ψ is the electron wave function in a quantum box, and $\rho \propto H$, the scaling effect always vanishes. By contrast, in three-terminal geometry, the flux across a quantum box can be described as M^2N, where $N = (I/e)\tau$ is the total population number of resonant states in energy window $E_0 < E < eV$, and the matrix element M couples ψ with the collector. It now becomes clear why only single I_c-V steps are manifested in the transmission spectra. As the opening of additional resonant channels does not change N, the higher order resonant states in these spectra are not visible.

Thus, in the quantum mechanical picture, a three-terminal experiment with a quantum box allows direct measurement of $|\psi|^2$. It seems remarkable that such measurement is facilitated by electron-phonon interaction, which represents the main mechanism of bulk relaxation in our experiment.

In conclusion, confinement-enhanced electron transport across an epitaxial metal-semiconductor interface was observed with the combination of STM and BEEM. Using modern methods of nanofabrication and *in situ* characterization, we have designed a three-terminal experiment with a *quantum box* and experimentally observed quantum scaling of its transmittance. Utilization of a quantum box in a three-terminal geometry may possibly open a way to development of nanoscale quantum transistors.

We thank K. A. Matveev, D. M. Chen, I. Shalish, and R. E. Martinez for fruitful discussions. We acknowledge support from the NSF through Harvard MRSEC (DMR 98-09363) and DARPA (DAAD 19-99-1-0158).

[1] M. Jalochowski, M. Hoffman, and E. Bauer, Phys. Rev. Lett. **76**, 4227 (1996).
[2] X. Tong, K. Horikoshi, and S. Hasegawa, Phys. Rev. B **60**, 5653 (1999).
[3] D. A. Evans *et al.*, Phys. Rev. Lett. **70**, 3483 (1993).
[4] I. B. Altfeder, K. A. Matveev, and D. M. Chen, Phys. Rev. Lett. **78**, 2815 (1997).
[5] J. J. Paggel, T. Miller, and T. C. Chiang, Science **283**, 1709 (1999).
[6] H. Hong *et al.*, Phys. Rev. Lett. **68**, 507 (1992).
[7] I. B. Altfeder, D. M. Chen, and K. A. Matveev, Phys. Rev. Lett. **80**, 4895 (1998).
[8] G. Binnig *et al.*, Phys. Rev. Lett. **50**, 120 (1983).
[9] W. J. Kaiser and L. D. Bell, Phys. Rev. Lett. **60**, 1406 (1988).
[10] M. Prietsch, Phys. Rep. **253**, 164 (1995).
[11] L. D. Bell and W. J. Kaiser, Annu. Rev. Mater. Sci. **26**, 189 (1996).
[12] L. D. Bell and V. Narayanamurti, Curr. Opin. Solid State Mater. Sci. **3**, 38 (1998).
[13] M. Prietsch and R. Ludeke, Phys. Rev. Lett. **66**, 2511 (1991).
[14] T. Sajoto *et al.*, Phys. Rev. Lett. **74**, 3427 (1995).
[15] M. E. Rubin *et al.*, Phys. Rev. Lett. **77**, 5268 (1996).
[16] T. Meyer *et al.*, Phys. Rev. Lett. **85**, 1520 (2000).
[17] M. Jalochowski and E. Bauer, J. Appl. Phys. **63**, 4502 (1988).
[18] J. A. Carlisle, T. Miller, and T.-C. Chiang, Phys. Rev. B **45**, 3400 (1992).
[19] Fitting the data to well-known BEEM theories [10,11] yields $E_0 = 0.83$ eV.
[20] A signal enhancement at the edges of the Pb islands and at the surface steps in the BEEM image of Fig. 2 is due to transient effects in the STM feedback loop.
[21] M. K. Weilmeier, W. H. Rippard, and R. A. Buhrmann, Phys. Rev. B **59**, R2521 (1999).
[22] R. C. Jaklevic *et al.*, Phys. Rev. Lett. **26**, 88 (1971).
[23] L. D. Bell, Phys. Rev. Lett. **77**, 3893 (1996).
[24] J. R. Anderson and A. V. Gold, Phys. Rev. **139**, A1459 (1965).
[25] J. Bardeen, Phys. Rev. Lett. **6**, 57 (1961).

BEEM IMAGING AND SPECTROSCOPY OF BURIED STRUCTURES IN SEMICONDUCTORS

V. NARAYANAMURTI, M. KOZHEVNIKOV

Gordon Mckay Laboratory of Applied Science, Harvard University, 9 Oxford Street, Cambridge, Massachusetts 02138, USA

ELSEVIER

AMSTERDAM - LONDON - NEW YORK - OXFORD - PARIS - SHANNON - TOKYO

ELSEVIER

Physics Reports 349 (2001) 447–514

PHYSICS REPORTS

www.elsevier.com/locate/physrep

BEEM imaging and spectroscopy of buried structures in semiconductors

V. Narayanamurti*, M. Kozhevnikov

Gordon Mckay Laboratory of Applied Science, Harvard University, 9 Oxford Street, Cambridge, MA, 02138, USA

Received October 2000; editor: D.D. Awschalom

Contents

* Corresponding author. Tel.: 617-495-5829; fax: 617-496-5264.
E-mail addresses: venky@deas.harvard.edu (V. Narayanamurti), mkozhevn@deas.harvard.edu (M. Kozhevnikov).

0370-1573/01/$ - see front matter
PII: S0370-1573(00)00119-8

Abstract

Ballistic Electron Emission Microscopy (BEEM) has been shown to be a powerful tool for nanometer-scale characterization of the spatial and electronic properties of semiconductor structures. In this article, we will discuss general aspects of BEEM experiment and theory in true ballistic and quasi-ballistic hot carrier transport. We will review the current state and recent progress in the use of the BEEM imaging and spectroscopy to study metal-semiconductor and metal-insulator-semiconductor interfaces, buried semiconductor heterojunctions and novel quantum objects. Various theoretical BEEM models are discussed, and their ability to describe BEEM experiments is examined. Special attention is drawn to the role of the electron scattering in the metal base layer, at the metal–semiconductor interface and in the semiconductor heterostructure on BEEM spectra.

PACS: 72.10.Fk; 73.20.At; 73.23.Ad; 73.40.Ns

Keywords: Ballistic transport; Tunneling microscopy; Semiconductor heterostructures

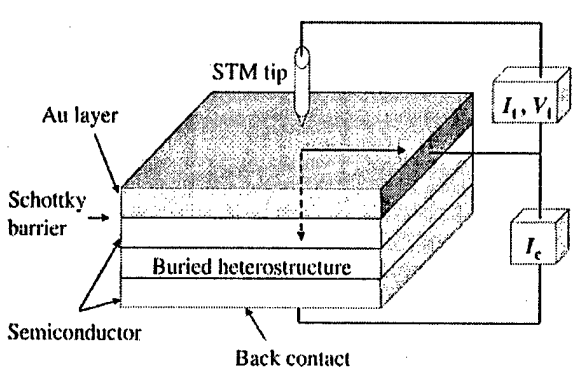
1. Introduction

The current trend of developing small-size electronic devices requires high resolution in the characterization and control of metal–semiconductor (m–s) heterostructures. To improve device characterization and performance, and for a better understanding of the basic physics of semiconductor interfaces, there is an increasing demand for nanometer-scale probes, all the way down to the atomic level, of the spatial and electronic properties of m–s junctions and semiconductor heterostructures.

Among the variety of scanning probe microscopy techniques, Ballistic Electron Emission Microscopy (BEEM), a three-terminal modification of Scanning Tunneling Microscopy (STM), is a powerful low-energy tool for nondestructive local characterization of semiconductor heterostructures. The STM metal tip (emitter) injects electrons across the tunneling gap into the metal (base) layer deposited on a semiconductor. A third terminal on the sample back is used to collect those electrons, which traverse the interface. A schematic of the BEEM experimental set-up is shown in Fig. 1. In the conventional STM study of semiconductors, the tunneling is usually between the metal tip and the doped semiconductor layers. In contrast, the BEEM technique provides the carrier filtration at the m–s interface, since only the electrons that can traverse the metal base and overcome the Schottky barrier will be collected at the semiconductor substrate (collector). A schematic band diagram for BEEM is shown in Fig. 2. In this way, the energy and angular distribution of hot electrons can be controlled independently on the semiconductor structure by simply varying the tip potential. In addition, the BEEM technique holds the extremely high spatial resolution of the scanning tunneling microscopy (~few nm). Thus BEEM provides, in complement to the surface morphology, a combination of low-energy electron microscopy and spectroscopy with high-spatial and energy resolution.

Since the pioneering work of Kaiser and Bell [1], applications of BEEM to various semiconductor surfaces and interfaces have already produced many interesting results (early BEEM research was previously reviewed in Refs. [2,3]). It was shown that the magnitude of the transmitted current into the substrate depends strongly on the local properties of the interface [4,5] as well as the scattering properties of the overlying metal film [6,7]. In epitaxial $CoSi_2$/Si structures, atomic BEEM resolution was observed [8], and this effect of electron beam focusing was attributed to be due to the silicide band structure [9]. The capability of BEEM to probe the electronic properties of semiconductors on the local scale has been demonstrated in several systems including InAs/GaAs [10], Si p–n junction [11] and SiGe strained layers [12]. It was shown that the high-bias BEEM spectroscopy could be used as a local density-of-states (DOS) spectroscopy in semiconductors [13,14]. By using BEEM, Schottky barrier and band structure characterization was conducted in technologically important semiconductors such as GaP [7], GaInP [15,16], GaN [17–19], GaAsN [20] and SiC [21,22].

Although the BEEM technique was originally invented as a unique microscopic and spectroscopic method to probe the Schottky barriers on a local scale [23], it has been successfully used to study the electronic properties of *buried* heterojunctions [24]. BEEM was successfully applied to metal–insulator–semiconductor (MIS) structures to study the transport related oxide properties on a microscopic scale in buried CaF_2/Si [25] and SiO_2/Si [26]. In SiO_2-based metal-oxide–semiconductor (MOS) structures, BEEM was used to study the quantum interference effects and trapped charge distribution that is very important for the operation of MOS field-effect transistors



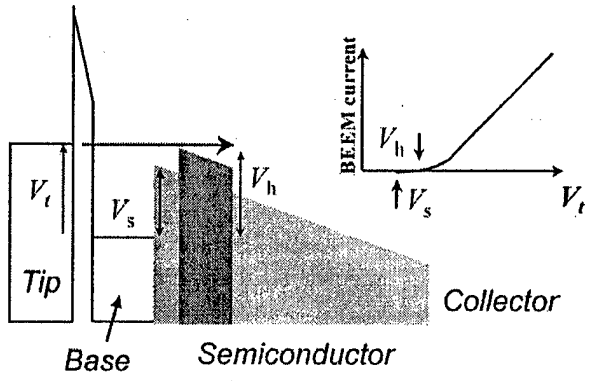


Fig. 1. A schematic of BEEM experimental setup.

Fig. 2. Schematic band diagram of a BEEM experiment. Inset shows a generic BEEM spectrum.

[27,28]. Recently, significant progress was accomplished in establishing BEEM as an effective method for measuring semiconductor heterojunction offsets [29–31], for measuring resonant transport through single barrier [32], double-barrier [33] and superlattice resonant tunneling heterostructures [34,35], for investigation of hot carrier transport in low-dimensional nanostructures such as quantum wires [36,37] and quantum dots [38–40], as well as for imaging of defects buried below the surface [41–43]. In order to amplify the BEEM capabilities, some apparatuses are operated in ultra high vacuum (UHV) [8,25,44] (the main advantages are cleaning of the semiconductor surface, metal deposition under UHV conditions, Schottky contacts and STM tips by using the metals and semiconductors that are unstable towards oxide formation) and at low (helium) temperatures [45,46] (the main advantages are enhanced energy resolution of BEEM and reduction of thermal noise and lateral drift).

2. BEEM transport

2.1. General BEEM description

The spectral shape of the BEEM current has to be known in order to derive a Schottky barrier height or heterostructure band alignment. A number of the theoretical models have been proposed to describe the BEEM spectrum in the threshold region. Two commonly used models, based on a planar tunneling formalism [47] and on transverse momentum conservation at the m–s interface, are the Bell-Kaiser (BK) model [1] and the Ludeke-Prietsch (LP) model [48]. In the BK model, for parabolic energy bands in the semiconductor, the near-threshold BEEM current varies as $I_c \sim (V_t - V_s)^2$, where V_t is the applied tip-to-base voltage and V_s is the Schottky barrier magnitude [2]. This model was found to adequately describe the BEEM spectra for the Au/Si system [1]. In the LP model, the theory is extended to include the energy-dependent electron mfp in the metal base layer and quantum mechanical transmission (QMT) at the m–s interface. The near-threshold dependence $I_c \sim (V_t - V_s)^{5/2}$ in this model differs somewhat from the BK model

result. It was reported in Ref. [48] that the best fit for metal/GaP BEEM data was obtained with the 5/2-power law. Also, the BEEM spectra of Au/GaAs were fitted by the 5/2-power law [49]. However, even though the BEEM correction due to the QMT is important, the quantitative difference between the BK and the LP models is comparable with the experimental error, and both of them can frequently fit the experimental data reasonably well [2,29,50,51]. Note that the assumption about $I_c \sim (V_t - V_s)^2$ (or $I_c \sim (V_t - V_s)^{5/2}$ in the case of the square-root contribution of the QMT) dependence is valid only close to threshold, while increasing deviations are expected for higher tip-to-base voltages. The main mechanisms responsible for these deviations are a voltage-dependent energetic distribution of the tunneling current, carrier scattering in the metal overlayer as well as carrier scattering and impact ionization in the semiconductor substrate [2].

A method for computation of the quantum mechanical transmission probability (QMTP) from plane interfaces was presented in Ref. [52]. This method was applied to calculate the QMTP at the metal/Si(100) and metal/Si(111) interfaces at the Schottky barrier maximum. It was shown that the QMTP depends on the crystal orientation and the location of the band maximum. Later, a more accurate BEEM model was developed [53,54] and was compared with the experimental results for Au/Si(100) [55] and Au/Si(111) [56,57]. The model employs an effective mass approach generalized to permit for non-diagonal mass tensors and for indirect conduction band minima. The impact of this quantum mechanical modeling on the BEEM current has been probed by Monte Carlo simulations (adopting the course developed in Ref. [56]) using bulk material parameters for all layers, and an average reduction of the BEEM current of 20% was obtained.

Recently, an extension of the BEEM theory for the case of buried heterostructures, was performed by Smith and Kogan (SK) [58]. In the SK model, the authors considered the QMT at the buried heterojunction interface in addition to the m–s interface. In addition, the authors have shown that the second voltage derivative of the BEEM spectra should reflect the transmission probability of the semiconductor heterojunction, and, therefore, is very informative for the characterization of the structure. For the case of the GaAs/AlGaAs systems [5,33], the experimentally obtained second derivatives (SD) of the BEEM current give peaked structure, which could be associated with the various calculated heterostructure transmission channels.

In the generalized BEEM description, four consecutive processes are usually considered: (i) tunneling from the tip to the metal base layer, (ii) hot electron transport through the metal base layer, (iii) transmission across the m–s interface (carriers must have enough energy and appropriate momentum in order to cross the Schottky barrier and reach the semiconductor), and (iv) transport inside the semiconductor that can include the transmission across the heterojunction interfaces. For the case of a buried heterostructure, only those electrons that pass through the m–s interface and are also transmitted by the heterostructure contribute to the collector current. In this case, the collector current can be written as a product of the electron flux distribution right before the semiconductor and the transmission coefficients of the m–s interface and heterostructure [59]:

$$I_c = eA \sum_{k^t} F_\perp^t(k^t) T^{V_t}(k^t, k^b) D(k^b) T^i(k^b, k^{sb}) T^h(k^{sb}) , \tag{1}$$

where A is the sample area, e is the electron charge, $F_\perp^t$ is the surface normal electron flux in the tip due to the state labeled by wavevector k^t, $T^{V_t}(k^t, k^b)$ is the voltage-dependent transmission coefficient for the state in the STM tip with wavevector k^t to tunnel through the vacuum into the base

state with wavevector k^b (determined from k^t by energy and surface normal wavevector conservation) calculated using the WKB approximation. $D(k^b) = e^{-\delta_b/\xi}$ is the probability of the electron traversing the base without scattering, where δ_b is the thickness of the base and ξ is the mean free path in the base. The product $F^t_\perp(k^t)T^V(k^t, k^b)D(k^b)$ is the interface normal electron flux incident on the m–s interface due to the state in the STM tip labeled by k^t. $T^i(k^b, k^{sb})$ is the transmission coefficient for the m–s interface and $T^h(k^{sb})$ is the transmission coefficient for the semiconductor heterostructure.

Note that, since there is a certain electron filtration for each of the aforementioned four steps, the collected BEEM current is typically smaller than the tunneling current by several orders of magnitude, depending on the structure under study. Also, applying forward tip-to-base bias (the tip is more positive than the metal base layer), electron tunneling is possible from the base to the tip, creating the ballistic hole distribution in the base. Therefore, allowing the collector to be a p-type semiconductor substrate, one can collect the above-threshold hole current in a manner quite similar to the case of electron injection. In this way, the BEEM theory can be extended for the case of hole injection, with a slight difference that the energy distribution of ballistic holes are peaked towards the base Fermi level. The details of the ballistic hole emission microscopy (BHEM) and spectroscopy (BHES) can be found, for example, in Refs. [60,61].

2.2. Impact ionization effect in semiconductors

Impact ionization in the semiconductor can contribute to the BEEM current at high voltages if energy loss in the metal film is weak so that the electrons reach the interface with the energy above the threshold for impact ionization in the semiconductor [62]. An additional onset of impact-ionized electrons in the BEEM current was observed at $V_t \sim 3V$ in Au/Si [51] and $CoSi_2$/Si [63]. For Au/GaP [7], this effect was found to be important in describing the BEEM data for 12 Å Au film whereas for 50 Å film, impact ionization was not important, due to strong inelastic scattering of hot electrons in the metal layer at high voltages.

The effect of impact ionization was studied in $NiSi_2$/Si(111)7 × 7 over a wide energy range (up to $V_t \sim 8V$) [44,64]. To eliminate the effect of elastic and inelastic scattering in the metal film, the BEEM spectra were taken in pinholes of thin $NiSi_2$ films where the modified Si(111)7 × 7 surface being an ultrathin metallic layer. Fig. 3 shows a representative BEEM spectrum taken in a pinhole of a 22 Å thick B-type $NiSi_2$ film. The spectrum is decomposed into a primary- and a secondary-electron part resulting from impact ionization. The quantum yield of impact ionization in Si (curve (c) in Fig. 3) was determined by decomposition of the BEEM current into secondary- and primary-electron contributions, and a good agreement with Monte Carlo simulations [65] was found. In addition, it was found that the secondary electron current starts at $eV_t \approx 1.5E_g$, where E_g is the semiconductor bandgap, in accordance with the theoretical predictions [66].

2.3. Image potential effect

The effect of phonon scattering in the semiconductor on the ballistic transport across the Schottky barrier was studied theoretically [67] and experimentally [55,68]. The authors distinguished between the metallurgical m–s interface and the Schottky barrier maximum resulting from the image potential effect. In this way, the Schottky barrier is not spatially abrupt, but exhibits

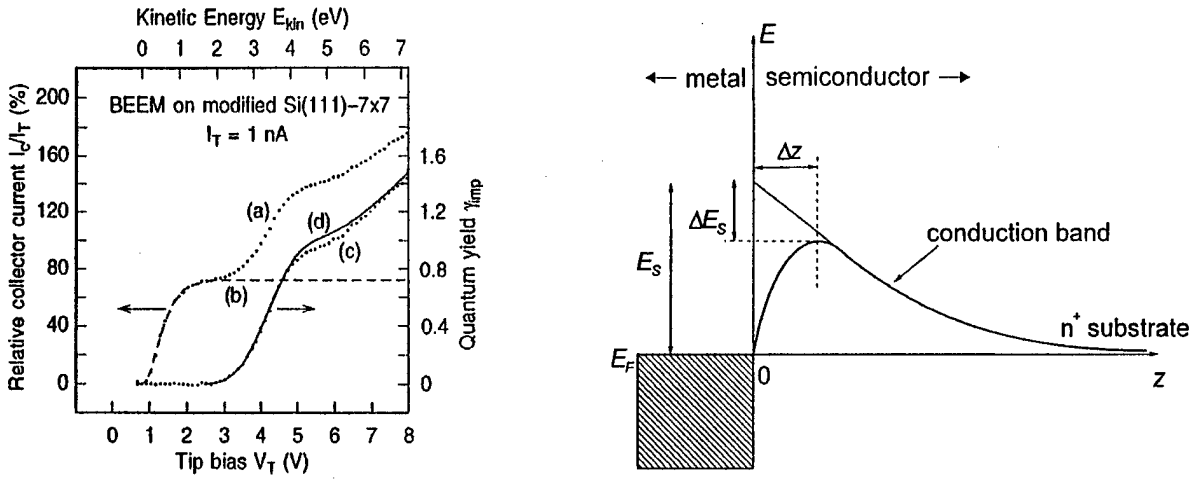


Fig. 3. BEEM spectrum for a 22 Å B-type $NiSi_2$ layer on Si(1 1 1)7 × 7 taken within a pinhole, curve (a). The dashed curve (b) is a fitted spectrum of the primary electron current. The difference (a)–(b) represents the secondary electron current arising from electron–hole pairs generated through impact ionization. This difference divided by (b), shown as the dotted curve (c), is the experimentally determined quantum yield γ_{imp}. The solid curve (d) is a calculation based on the model in Ref. [66] (from Ref. [64]).

Fig. 4. The effect of the image potential on the Schottky barrier lowering and on the shift of the potential maximum of the barrier.

a smooth hydrogen-like variation (see Fig. 4). In the case of band bending, this results in a lowering of the effective barrier height by $\Delta E_s \approx \sqrt[4]{e^6 N_D E_s/(8\pi^2(\varepsilon_s\varepsilon_0)^3)}$, where N_D is the doping concentration of the semiconductor, and in the barrier maximum shifting away from the interface by $\Delta z \approx \sqrt[4]{e^2/(2^{10}\pi^2\varepsilon^s\varepsilon N E_s)}$ [2]. Electrons with energies just over the threshold for transmission that excite phonons in the region before the Schottky-barrier maximum are expected to have a high probability of reentering the metal [55]. It was shown that while QMTP alone adds $\sim E^{1/2}$ dependence, the combined effect of optical phonon scattering and quantum mechanical scattering gives a much weaker energy dependence than for either effect alone for the cases of Au/Si and Au/GaAs at both $T = 77$ and 300 K [67].

While the image potential effect on the effective Au/GaAs Schottky barrier is mostly insignificant, its effect on the barrier maximum shift has much more important consequences. Indeed, assuming $N_D = 5 \times 10^{15}$ cm^{-3}, $\varepsilon_s = 12$, $E_s = 0.95$ eV at $T = 300$ K for a typical Au/GaAs Schottky contact, we obtain $\Delta E \approx 20$ meV and $\Delta z \approx 30$ Å. ΔE_s of 20 meV will not noticeably affect our measurements since it is only $\sim 2\%$ of the Schottky barrier height and is comparable with the experimental error of ~ 30 meV. On the other hand, Δz of 30 Å is comparable, and even longer, than the expected mfp length for the X-electrons (~ 10 Å at $T = 300$ K [43]), and, therefore, can affect this transport channel very effectively. It was confirmed in our BEEM study of the Au/GaAs system, where a strong attenuation of the electron transport through the X valley was observed [5].

The effect of phonon scattering beyond the Schottky-barrier maximum on the magnitude of the BEEM current depends on the build-in electric field strength, i.e. on the doping density of the

semiconductor. The effect of the image potential on the Schottky barrier reduction under reverse bias applied was studied in Au/Si [68]. The increase of the collection efficiency under applied reverse-bias ($\sim$10% increase of the collection efficiency at a 3 V reverse-bias voltage) was explained by a reduction of the energy-dependent backscattering of electrons in the depletion layer. In another study, it was also shown that the electron backscattering into the metal due to the interactions with phonons inside the depletion layer near the interface is very important for an adequate description of the BEEM spectra in the Au/Si system [69].

2.4. Carrier scattering at the metal–semiconductor interface

While the original model has been subsequently generalized to include effects such as quantum mechanical transmissions at the m–s interface and in buried semiconductor heterostructures, the assumption of transverse wavevector conservation at the m–s interface has been mostly maintained. Using such models good fits to threshold data have been achieved for a number of metal/semiconductor systems. However, even if the assumption of transverse wavevector conservation at the metal-semiconductor interface may be optimistically justified at epitaxial interfaces such as $CoSi_2$/Si and $NiSi_2$/Si, it is very questionable for the case of nonepitaxial m–s interfaces where the interface is far from the atomically abrupt, such as Au/GaAs or Au/Si that are prototypes for nonepitaxial metal interfaces on direct and indirect bandgap semiconductors, respectively. In this case, the carrier transport across the m–s interfaces may be quasi-diffusive rather than ballistic due to the carrier scattering at these interfaces.

2.4.1. Nonepitaxial Au/Si and Au/GaAs

Tunneling predominantly injects forward-directed electrons with small interface parallel wavevector components into the metal film. As shown in Fig. 5, for Si(111), for which all of the CBM have large parallel wavevectors, there are no states available for transmission at the center of the interface Brillouin zone (BZ), while for Si(100), for which two of the conduction valleys have zero parallel wavevector, there are states available both at the zone center and in off-axis band minima. Thus, the Au/Si(111) BEEM current is expected to be much weaker than that for Au/Si(100) if the electron transport is essentially ballistic through the m–s interface [59]. This predicted difference was not confirmed in the BEEM experiments. A deviation from the ballistic picture was experimentally observed for Au/Si [1,70,71], and Pd/Si [13] systems, where essentially the same BEEM spectra were observed for Si(001) and Si(111) substrates (see Fig. 6(a)), in conflict with conventional BK theory which predicts a significantly different spectra for the two orientations.

With the same argumentation for the case of Au/GaAs(100), the calculated current into valleys with zero parallel wavevector at the minimum (e.g., the Γ valley), is predicted to be much larger than the calculated current into valleys with large parallel wavevector at the minimum (e.g., the L valleys) [59]. In BEEM experiments on Au/GaAs, three thresholds were observed and were attributed to the electron transmission into Γ, L and X valleys [29,49]. Six fitting parameters were commonly used to describe the BEEM measurements in GaAs (AlGaAs), one threshold energy and one scaling factor for each of the three valleys. And though the comparison of the multiple BEEM thresholds with the expected composition dependence of the GaAs Γ, L, and X conduction band position yielded reasonably good agreement over the entire compositional range, the scaling factors differ significantly from the expected ones, as shown in Fig. 6(b). In fact, the Γ channel

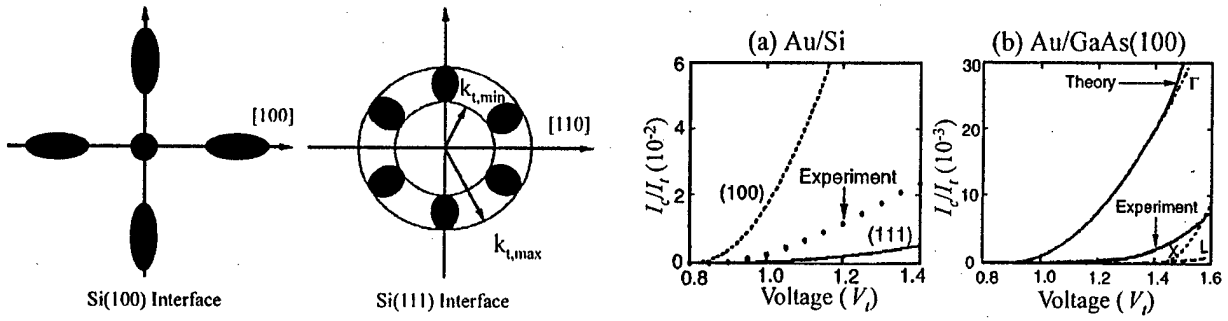


Fig. 5. Schematic representation in k space of the projection of the six Si CBM onto the (1 0 0) and (1 1 1) interfaces. The transverse wave vectors of the incident electrons must lie within the shaded regions in order to satisfy energy and transverse momentum conservation laws. It is clear from the diagram that for energies just above the CBM, electrons with near zero transverse momenta can cross a Si(1 0 0) interface but are reflected back into the metal at a Si(1 1 1) interface [72].

Fig. 6. (a) Comparison of the calculation for Au/Si(1 0 0) (dashed line) and Au/Si(1 1 1) (solid line) with the experimental results for Au/Si of Refs. [86,87]. These measurements were made on Au/Si(1 1 1) but measurements for Au/Si(1 0 0) are very similar to those for Au/Si(1 1 1) (see Refs. [70,102]). (b) Comparison of the calculation for Au/GaAs(1 0 0) with the experimental results of Ref. [29]. The dot-dashed line shows the Γ, the long-dashed line shows the L, the short-dashed line shows the X contribution to the BEEM current, and the solid line shows the sum of the three [85].

contribution to the collector current was found to be much smaller than the L channel contribution [1,29,49]. The BEEM current into the L valley was explained in terms of possible breakdown of the transverse momentum conservation at the m–s interface [5,29,49]. In Au/GaAs/AlAs [49], two thresholds were associated with the contribution from the X and L valleys of AlAs. Complementary Monte Carlo simulations give a good agreement with the experiment if one assumes transverse momentum conservation at the GaAs/AlAs interface and its non-conservation at the Au/GaAs interface.

2.4.2. *Epitaxial $CoSi_2$/Si and $NiSi_2$/Si*

Epitaxial metallic disilicide-silicon films such as $NiSi_2$ and $CoSi_2$ on Si can be grown with atomically abrupt and structurally perfect interfaces either by the Tung template method [72,73] or by stoichiometric coevaporation of Co and Si at room temperature and subsequent annealing [74,75]. In these systems, the role of the transverse momentum conservation at the m–s interface was explicitly studied in a number of BEEM experiments. While the assumption of transverse momentum conservation was justified in some cases, a certain deviation from the behavior, expected if transverse momentum $k_{||}$ is conserved, was observed in other cases.

The BEEM current depends strongly not only on the Schottky barrier but also on the electronic band structure of the material under study. Based on their band structure calculations, Stiles and Hamann [76] have predicted a delayed onset of the BEEM current for the $CoSi_2$/Si(111) system where there are no states in the silicide with $k_{||}$ large enough to match the silicon CBM up to 0.85 eV above E_F (Schottky barrier for this system is ~ 0.66 eV). Fig. 7 shows schematically part of the distribution of states in both materials in terms of their energy and parallel wavevector in the

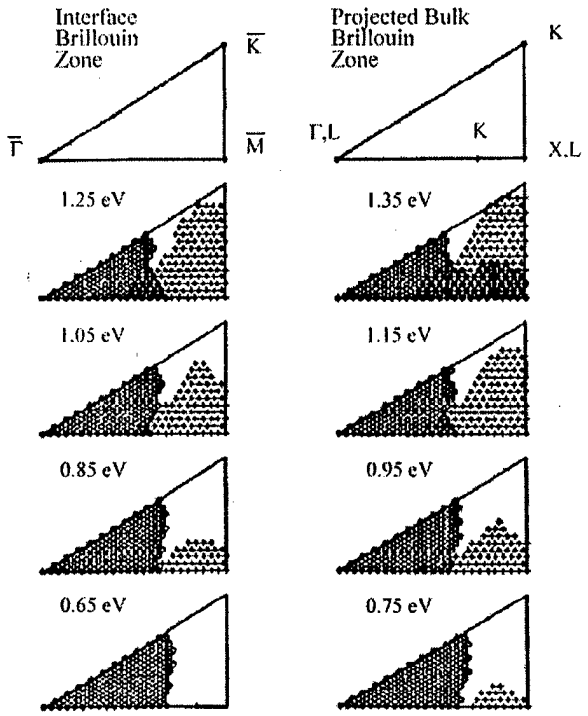


Fig. 7. Phase space for electron transmission through a $CoSi_2$/Si(1 1 1) interface. The panels show the irreducible wedge of the interface Brillouin zone of both $CoSi_2$ and the Si. At each parallel wave vector used in the calculation there is an open circle if there is at least one state at that energy in the $CoSi_2$ and a plus if there is at least one state in the Si. If there is a state in both, the open circles fill in and become closed circles [76].

interface BZ. In agreement with this prediction, Kaiser et al. [74] have measured the BEEM spectra in this system in a glove box and found a delayed onset of $\sim$0.85 eV, close to the calculated value. The observation of this effect was considered as a strong evidence for the transverse momentum conservation in the $CoSi_2$/Si(111) system.

Later, Sirringhaus et al. [75,77,78] observed individual interfacial dislocations by conducting the BEEM imaging and spectroscopy on strain-relaxed $CoSi_2$ layers. The dislocations were also observed in STM images because the strain field around each dislocation at the interface causes a distortion of the surface. The sharp BEEM dislocation profiles ($\sim$ 10 Å resolution for a 25 Å $CoSi_2$ film) in $CoSi_2$/Si(111) imply that the electron transport in the silicide is essentially ballistic. This was also corroborated by the observation of the quantum size effect [8,79]. Thus it was concluded that most of the electrons reach the $CoSi_2$/Si interface with $k_{||} \sim 0$. The increase of the BEEM current on the dislocation was explained by increased interfacial scattering at the dislocation core that supplies electrons with a large transverse momentum to transmit into the silicon CBM. This scattering broadens the $k_{||}$ momentum distribution, and thus facilitates the electron transmission. In addition, if $k_{||}$ is conserved away from dislocations, the electron scattering will not be beneficial in the case of positive biasing of $CoSi_2$/*p*-Si(111), where holes are injected into the zone-centered valence band. In accordance with expectations, in the latter case the BEEM current at dislocations and at isolated defects was lower than that in a defect-free region [80].

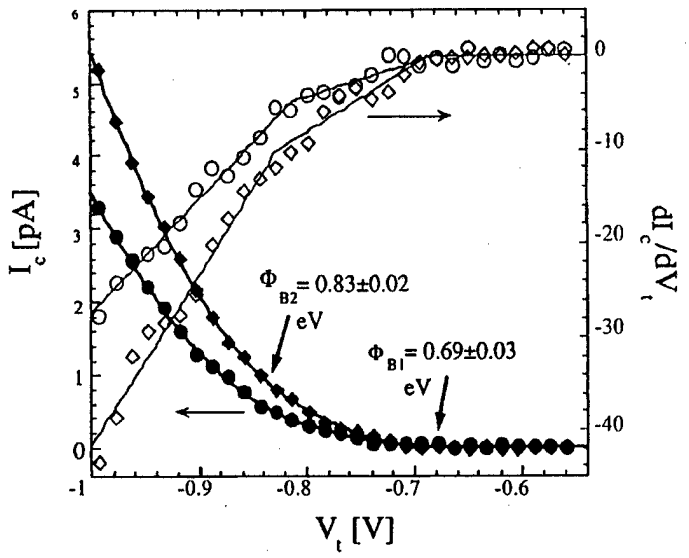


Fig. 8. Average of 14 BEEM spectra, and corresponding first derivative spectra, taken on top of an interfacial point defect (diamonds) and in a defect-free area (circles). The spectra were measured with a tunneling current of $I_t = 20$ nA, and are normalized to $I_t = 1$ nA [82].

However, in all in-situ measurements of Sirringhaus et al. [81], a homogeneous BEEM onset of ~0.66 eV was found (corresponding to the Schottky barrier) in dislocation- and defect-free regions, and it was explained by a small but finite scattering probability everywhere at the epitaxial $CoSi_2$/Si interface. Later, better-spatially resolved BEEM measurements (where I_t was increased to 20 nA to allow less averaging to obtain a high signal-to-noise ratio), revealed that both onsets (~0.69 and ~0.83 eV) should be invoked to fit the BEEM spectra [82], as shown in Fig. 8. One can see from Fig. 9 that the amplitude weight of the lower onset is highly increased in the presence of the interfacial defects, indicating an enhanced scattering probability. Also, in the case of epitaxial $CoSi_2$/Si(111) and $CoSi_2$/Si(100), similar BEEM current magnitude at the near threshold region was observed [41]. This discrepancy from the behavior expected if $k_{||}$ is conserved was explained by the combined effect of small interface transmission at the $CoSi_2$/Si(100) interface and the broad momentum distribution injected by the atomically sharp STM tip.

Another epitaxial system with an abrupt interface is $NiSi_2$/Si(111). In contrast to $CoSi_2$/Si(111), in this system there is overlap between $NiSi_2$ and Si states at all energies above the Schottky barrier, and transverse momentum conservation cannot be manifested just from a shifted threshold position. However, if transverse momentum is conserved, a very soft threshold is expected for electron transport through the X valley with large parallel wavevector, and, indeed, a soft threshold was found in Ref. [44]. Another manifestation of transverse momentum conservation in $NiSi_2$/Si(111) is the increase of BEEM current at the edges of the Si triangles on top of the $NiSi_2$ terraces [44]. In the presence of a step (at the edges of the islands) the injected electrons are dispersed, and the angular distribution is expected to be broader thus leading to a larger transmission. In another study of epitaxial $NiSi_2$/Si(111) system, however, the deviations of the transmittance spectra of $NiSi_2$/Si(111) from the theoretical predictions of Ref. [83] were observed

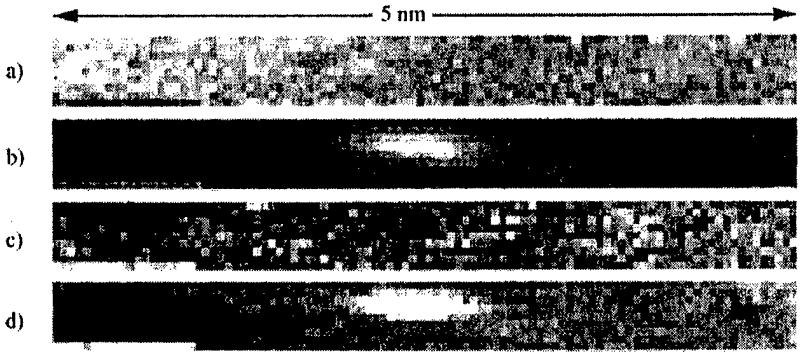


Fig. 9. Topographic STM image (a) and corresponding BEEM image (b) obtained for $V_t = 1.7$ V and $I_t = 20$ nA. A total of 100×9 BEEM spectra was recorded at a nominal spacing of 0.05 nm. The elongated shape of the point defect is due to microscope drift during the measurement. Each spectrum was fitted to $I_c(V_t) = R_1(V_t - \Phi_{B1})^2 + R_2(V_t - \Phi_{B2})^2$ keeping the barriers fixed at $\Phi_{B1} = 0.69$ eV and $\Phi_{B2} = 0.83$ eV. The scale factor R_2 deduced from the fit is shown in (c). Similarly, the scale factor R_1 is displayed in (d). Gray-scale ranges were 0.1 nm, $\Delta I_c = 200$ pA, $R_2 = 28$–84 pA/(eV)^2, and $R_1 = 14$–42 pA/(eV)^2 in (a), (b), (c) and (d), respectively [82].

and were attributed to the elastic scattering in the silicide layer [71]. It was found that while the cross-sectional TEM and X-ray scattering data show good epitaxial interfaces and silicide layers, a 10% scattering in the silicide should be assumed (in the isotropic point defect scattering model) to describe the experimental BEEM data. It was concluded that scattering broadens the momentum distribution and thus averages and smoothes the effect of band structure predicted by theory.

2.4.3. Scattering models

Several models were employed to describe the observed deviation from the ballistic picture in BEEM experiments on metal/Si and metal/GaAs systems.

Ke et al. [49] used a Monte Carlo technique to simulate the I–V curve of the BEEM measurements in Au/GaAs and Au/GaAs/AlAs. Four physical processes were taken into account: electron tunneling from the tip to metal surface, elastic and inelastic scattering in the metal base layer, electron transmission through the interface and impact ionization inside the semiconductor depletion layer. It was shown that the experimentally observed large BEEM current through the L valley can be explained only assuming strong scattering at the m–s interface. Good agreement between Monte Carlo simulations and experimental data was obtained by assuming the square-root energy dependence of the electron transmission at the interface for all valleys, regardless of their projection in the (001) direction.

Guthrie et al. [84] have developed a dynamic BEEM model, where in addition to directional (kinematic) constraints, the amplitude effects were incorporated by considering the quantum mechanical transmittance and elastic scattering in the base and at the m–s interface. It was shown that as the scattering probability increases (i.e. a planar distribution is replaced by an isotropic distribution) the Au/Si(111) BEEM current increases whereas the Au/Si(100) current decreases. In addition, the BEEM spectral shape is also sensitive to the scattering probability. The model gives a good fit to the experimental data for both Au/Si(111) and Au/Si(100) if the scattering probability at the m–s interface exceeds 50%. Also, including the effect of an energy-dependent quantum

transmittance improves the fit in the near-threshold region. In the case of Au/GaAs(100), the BEEM current through the Γ valley decreases strongly as the scattering strengthens, without any spectral shape change. From fitting the experimental data, it was difficult to determine the extent of the electron scattering since the BEEM spectral shape is not sensitive to the scattering, and the scaling factor can be strongly modified by the surface passivation and the effect of electron scattering in the metal overlayer.

To analyze quantitatively the experimental data, the m–s interface-induced scattering (MSIS) model was proposed in Ref. [85]. The authors showed that the experimental data for Au/Si and Au/GaAs systems could be fitted only by taking into account of the strong carrier scattering at the m–s interface. Recently, to quantitatively analyze the electron scattering at the m–s interface, this model was modified to include anisotropy of the electron effective mass, the energy dependence of the electron mfp in the metal base, and finite temperature [59]. In this model, the conduction process is described by carrier multivalley transport with corrections due to the scattering at the m–s interface, which depends on the m–s scattering probability. When interface scattering is considered, the BEEM current is given by [59]

$$I_c = eA\sum_{k'} F_{\perp}^{t}(k^t)T^{V_t}(k^t,k^b)D(k^b) \times\left[\left(1-\sum_{f} P_{k^b,k^f}\right)T^{i}(k^b,k^{sb})T^{h}(k^{sb})+\sum_{f} P_{k^b,k^f}T^{i}(k^b,k^{sb})T^{h}(k^{sb})\right], \tag{2}$$

where P_{k^b,k^f} is the probability for the electron with the wavevector k^b to be scattered into the state labeled by the wavevector k^f. The first and second terms in brackets correspond to electrons that are not scattered and that are scattered, respectively. In the weak scattering limit, the MSIS model is essentially equivalent to the SK model.

In this model, if scattering at the m–s interface is strong, the injected electron flux is redistributed, and valleys with zero interface transverse wavevector at their energy minimum are not preferentially weighted. Instead the weighting goes like the density of final states for the scattering process. Fig. 10 compares the calculated BEEM current in the strong and in the weak scattering limits for Au/Si with the experimental results of Refs. [86,87]. One can see that the experimental data can be described adequately without fitting parameters only in the strong scattering limit. In BEEM on Au/GaAs, the calculated Γ channel contribution is much larger than the L channel contribution to the collector current in the weak scattering limit, whereas the L channel contribution is much larger than the Γ channel contribution in the strong scattering limit (see Fig. 11). Scaled threshold shapes are similar in the weak and strong scattering limits for conduction band valleys that make a circular projection on the interface plane, such as Γ and X valleys. For conduction band valleys that make a strongly anisotropic projection on the interface plane, such as the L valleys of GaAs(100), the calculated L channel currents have significantly different threshold shapes in the two limits, as shown in Fig. 12. When the calculations are done in a spherical approximation, the threshold shapes are similar in the two limits and look like the result of the more complete anisotropic model in the strong scattering limit.

Recently, the electron scattering at the m–s interface was studied experimentally in the Au/GaAs/$Al_xGa_{1-x}As$ system by using SD-BEEM spectroscopy [5]. The SD-BEEM spectrum is proportional to the heterostructure transmission coefficient [58], and is particularly useful both for

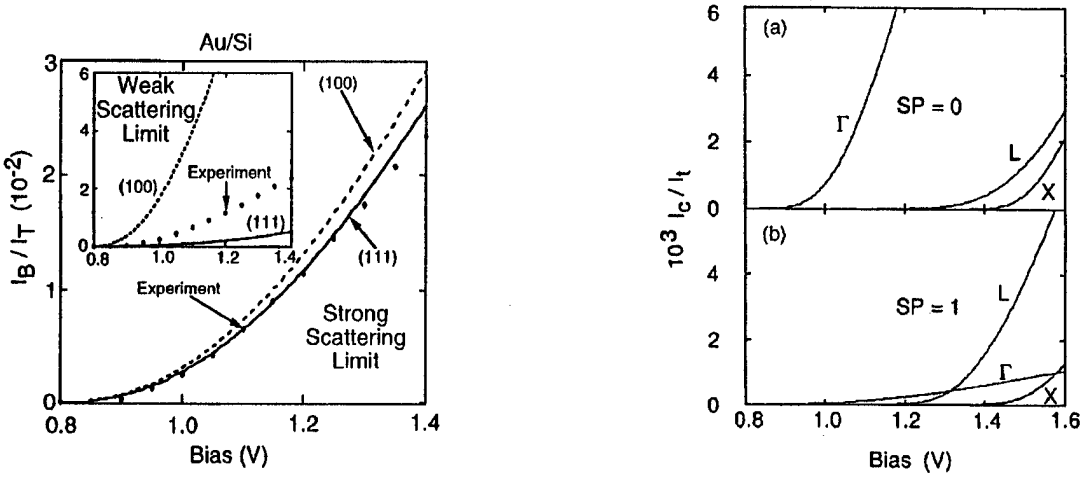


Fig. 10. Comparison of the calculation in the strong scattering limit for Au/Si(1 0 0) (dashed line) and Au/Si(1 1 1) (solid line) with the experimental results for Au/Si of Refs. [86,87]. These measurements were made on Au/Si(1 1 1) but measurements for Au/Si(1 0 0) are very similar to those for Au/Si(1 1 1) (see Refs. [70,102]). The inset (referenced earlier in our paper as Fig. 6(a)) compares the calculation in the weak scattering limit to the same experimental results [85].

Fig. 11. Calculated ratio of collector to tunneling current as a function of bias for the three conduction band channels of Au/GaAs(1 0 0) in the weak scattering limit (upper panel) and in the strong scattering limit (lower panel) [59].

separating the contributions from various conduction band channels and for analyzing the threshold shapes. The analysis of the SD-BEEM spectral shape is more effective than just the BEEM current magnitude because the absolute value of the BEEM current could be scaled by the fluctuations in both the metal layer and semiconductor epilayer thickness. In GaAs, conduction bands with both zero and large interface parallel wavevector coexist in the same material. Thus direct and indirect transport channels can be investigated in the same sample. The observed SD-BEEM spectral shape shows obvious deviation from the calculations that assume transverse momentum conservation. In Fig. 13, we compare the calculated (dashed lines) collector current and its second voltage derivatives with the corresponding measured quantities (solid lines) for a Au/GaAs sample at room temperature. (Sample structures and the measurement technique are described in Ref. [5].) The SD-BEEM spectrum was extracted from the experimental BEEM spectrum in Fig. 13(a) by numerical differentiation with a 10 meV window, as shown in Fig. 13(b). The probability of electron scattering at the m–s interface (SP) was adjusted to fit the SD-BEEM spectra. The SD-BEEM spectra clearly separate the contributions from Γ and L electrons; thus, the weighting of the relative Γ and L channel contribution is a sensitive test of the model. The best fit to the SD-BEEM spectrum gives a 85% probability for electron scattering at the m–s interface. For comparison, theoretical curves for the cases without scattering and with 100% scattering are also shown. The measured relative contribution of the L-electrons is different than that calculated for the L channel in the weak scattering limit. Specifically, the L-electron collector current is found to be the strongest one, whereas, assuming transverse momentum conservation, the contribution of

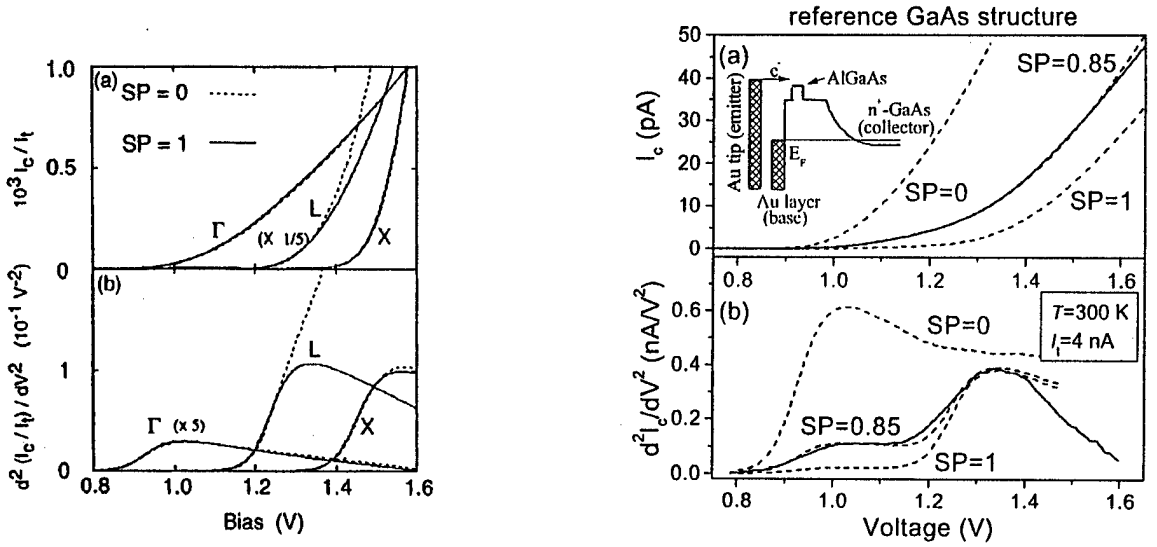


Fig. 12. Scaled ratio of collector to tunneling current (upper panel) and scaled second voltage derivative of the ratio of collector to tunneling current (lower panel) as a function of bias for the three conduction band channels of Au/GaAs(1 0 0). The scaling factors for all the valleys were chosen so that the currents for two limits overlap well in the near-threshold region. The strong scattering limit results (solid lines) are as calculated and the weak scattering limit results (dashed lines) are scaled by: Γ channel, divide by 31; L channel, multiply by 4.5; X channel, divide by 1.7 [59].

Fig. 13. The room-temperature BEEM (a) and SD-BEEM (b) spectra of the 1 μm undoped GaAs layer grown on n^+ GaAs substrate. The MSIS model calculations (dashed lines) are also presented for three values of SP, the electron scattering probability at the m–s interface. The inset in (a) shows a schematic band diagram of the samples under study [5].

the L band (which does not project to the zone center of the interface BZ) can give only a small BEEM current onset. The Γ and L channel contributions are included in the calculation but the X channel contribution is not included. In this experimental structure, the absence of the X conduction minimum contribution at room temperature is due to the image potential-induced strong electron scattering in the spacer between the metallurgical m–s interface and the maximum of the barrier height [2,67]. Fig. 14 presents the MSIS model fits to the BEEM and SD-BEEM spectra for a number of Au/GaAs/$Al_xGa_{1-x}As$ single barrier structures. In this case, the best fits are obtained with the scattering probability at the m–s interface varying between 85% and 92% among these samples. This small variation in the scattering parameter indicates that our diode fabrication procedure is reproducible and results in approximately the same quality of the m–s interface. Intervalley scattering at the GaAs/AlGaAs heterojunction interfaces is assumed to be minor in comparison with the scattering at the m–s interface. Indeed, intervalley scattering in GaAs/AlGaAs heterojunctions was studied both experimentally [88,89] and theoretically [90,91], and conservation of the transverse momentum is expected for this epitaxial heterojunction interface. For epitaxial growth, the GaAs/$Al_xGa_{1-x}As$ interface is a near perfect heterojunction interface, with a lattice mismatch of less than 0.8%. In contrast, metal deposition is not epitaxial,

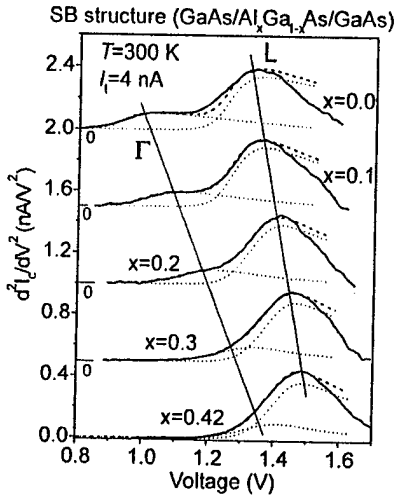


Fig. 14. Room-temperature SD-BEEM spectra for five different Al compositions (solid lines). For clarity, the SD-BEEM spectra are shifted along the vertical axis. Thin solid lines are eye-guides for the peaks position development. The MSIS model calculations are also presented. The model calculations show the separate Γ and L valley contributions (dotted lines) and their sum (dashed lines) [5].

and, therefore, the quality of the Au/GaAs interface is not as good as the quality of the GaAs/$Al_xGa_{1-x}As$ heterojunction interface. The possible effect of electron multiple reflection either in the metal (the metal layer is thick enough that a multiply reflected component is strongly attenuated) or at the heterostructure interfaces (this mechanism would amplify mainly the current of the Γ-electrons since they are the electrons with the longest mfp in the semiconductor heterostructure) is small and cannot explain the observed strong L-electron and comparatively weak Γ-electron contributions.

2.5. *Metal base layer effect on BEEM transport*

2.5.1. *Effect of carrier scattering in metals*

The BEEM current transmitted through the metal base layer attenuates strongly as the metal film thickness increases. The survey of the experimental and theoretical data is presented in Fig. 15. The total attenuation length λ in the metal can be described by $1/\lambda = 1/\lambda_I + 1/\lambda_e$, where λ_e is the electron elastic mfp and λ_I is the electron inelastic mfp in the metal. The energy dependence of λ is due to the energy dependence of λ_I that, at low voltages, results from inelastic electron–electron scattering processes [2,86]. λ_I is usually given in the form of $\lambda_I(E) = E/[\Lambda(E - E_F)^2]$, expected for the electron–electron scattering mechanism [69], where Λ is the electron–electron scattering strength parameter. Taking into account the Au band structure, Reuter et al. [69] studied the energy dependence of the inelastic mfp (due to the electron–electron interactions), using a Green-function approach and complementary Monte Carlo simulations. The authors found that λ_I predicted by the standard Fermi-liquid theory provides excellent agreement between theoretical and experimental BEEM spectra (giving $\lambda_I(E - E_F = 1\text{ eV}) = 252$ Å). In contrast, a relatively weak linear energy dependence in the empirical form of $\lambda_I(E) = 1.5\lambda_{i0}[1 - (E - E_F)/3]$ was deduced

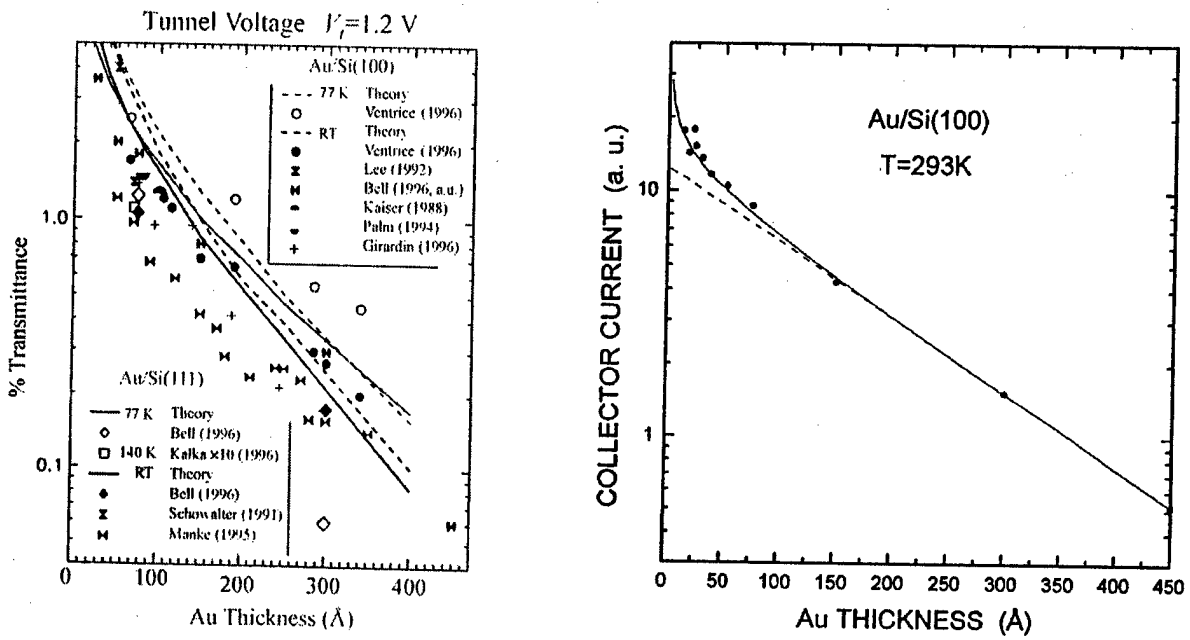


Fig. 15. Transmittance of the Au/Si interface at a tunnel voltage of 1.2 eV as a function of the Au-layer thickness [103].

Fig. 16. Experimental BEEM attenuation length measurement (circles) for Au/Si(100) samples. I_c was measured at $V_t = 1.2$ V. The solid (dashed) line is the dependence, which results if boundary reflections within the Au film are allowed (not allowed) [86].

from the fitting procedure for the Au/Si system [86], where λ_{i0} is the inelastic mfp length for electrons at $E - E_F = 1$ eV, using $\lambda_e = 400$ Å and $\lambda_{i0} = 220$ Å. The fitted total attenuation length of $\lambda \sim 130$ Å at 1 eV above the Fermi energy is in a good agreement with $\lambda = 133$ Å at $T = 300$ K and $\lambda = 147$ Å at $T = 77$ K obtained in Au overlayers on Si(100) over the energy region of 0.92–1.2 eV above the Fermi level [55], even though a certain distribution in λ values is also expected [92,93].

If λ_I is much shorter than λ_e, λ will also be energy dependent, namely, will decrease with the electron energy (tip-to-base voltage) increase. It will induce, according to Eq. (1), a certain additional decrease of the BEEM current as the tip-to-base voltage increases, therefore, changing the BEEM spectral shape. Comparison of Monte Carlo simulations with the experiment shows that the strong variations in the spectral shape of the BEEM current in metal/GaP at $V_t < 4$ V are mainly due to the energy dependence of the mfp for inelastic scattering in the metal. Ludeke et al. [64] showed for the case of Pd/GaP that, when the base layer is much thicker than the mfp of electrons, the multiple collisions with phonons will reduce effectively the injected electron energy and induce the delay of the onsets observed in the BEEM current. For thin metal layers, the effect of the electron scattering in the metal (Cr, Pd and $NiSi_2$ at thickness < 100 Å) was mainly the attenuation of the BEEM current without the spectral shape change. While the truly ballistic transport in the metal film is obscure in most experiments [7], the ballistic characteristic would strengthen if experiments were performed in ideal conditions using clean tips and samples at low

temperatures and low voltages. In fact, features due to standing waves in thin metal films were observed in experimental BEEM spectra [77].

By studying the dependence of BEEM spectral shape on Au thickness, Bell [86] argued that the contribution of multiple electron reflections within the metal layer is the primary factor preventing previous BEEM observations of $k_{||}$ conservation in thin metal films. To illustrate the effect of multiple electron reflections in the metal layer, the authors plotted the collector current dependence on the Au thickness in Au/Si(100) system, as shown in Fig. 16. The deviation from the logarithmic dependence in the limit of thin metal layers was attributed to the presence of multiple electron reflections. Recently, Reuter et al. [69] have used a more advanced model to fit the experimental data of Ref. [86], and the importance of the multiple electron reflection in describing the spectral shape and amplitude of the BEEM current was confirmed.

2.5.2. *Metal band structure effect*

Previously discussed models to explain the lack of BEEM sensitivity to the Si orientation in the Au/Si systems have usually invoked strong elastic scattering of the ballistic electrons, either at the Au/Si interface, and/or in the bulk or at the surface of the Au overlayers. In contrast, by using the Keldysh Green-function method, Garcia-Vidal et al. [94,95] suggested that the Au band structure might be responsible for the observed similarity of Au/Si(100) and Au/Si(111) BEEM spectra (at least both mechanisms, i.e. $k_{||}$ non-conservation and Au band structure effect, should be weighted). It was pointed out that the Au films often grow (111) oriented on Si surfaces after the first four or five layers are completed, and that there are no propagating Bloch electron states in the cone centered on the (111) direction, from below the Fermi surface to several eV above it. This requires that electrons injected into such an overlayer must propagate at angles considerably away from the normal if they are to travel ballistically through the metal to the Au/Si interface. This approach

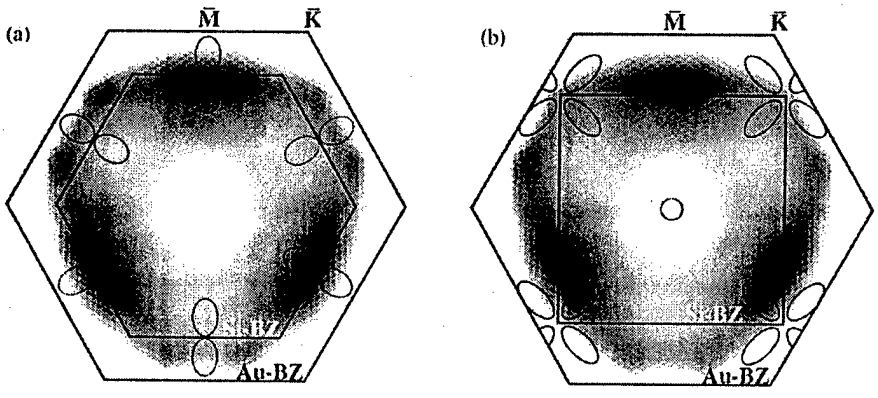


Fig. 17. Matching of the obtained semiclassical k-space BEEM current distribution with the available CBM states in the semiconductor, which are approximated by projected parabolic bands. Projected gold surface BZ is given by external hexagon, whereas Si(1 1 1) and Si(1 0 0) BZs are, respectively, given by the hexagon and the square inside. Current can only enter into the Si through the area enclosed by the ellipses ($E = E_F + 1$ eV). (a) Au/Si(1 1 1), $E_s = 0.86$ eV, (b) Au/Si(1 0 0), $E_s = 0.82$ eV. Note that the different lattice parameters of Au and Si require remapping of the Si ellipses inside the larger Au BZ. $\bar{\Gamma}$–$\bar{M}$ corresponds to the [1 0 1] direction in k space [95].

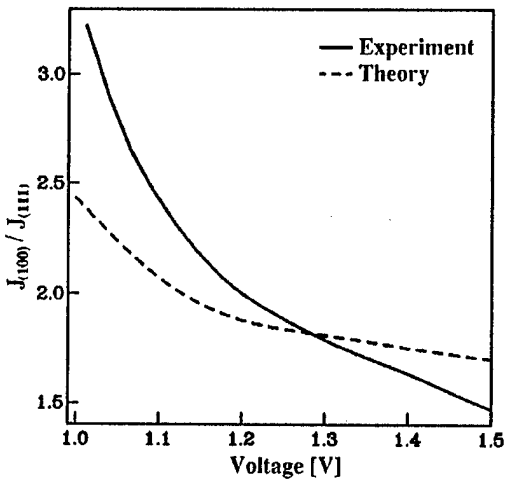


Fig. 18. Ratio of current injected in Si(1 0 0) and Si(1 1 1) after propagation through thin Au films of approximately equal thickness (100 Å Au/Si(1 0 0) and 75 Å Au/Si(1 1 1), chosen to compare with available experimental data). Elastic electron propagation is used in the theory (dashed line); experimental data taken from Refs. [86,96] (solid line) [95].

differs from the free-electron treatment of the existing model [1] that allows an electron beam propagating in normal direction through the metal film grown in the [111] direction on either Si(111) or Si(100). For Au(111) on Si(100) and Si(111), the allowed (in the metal) wavevector and the silicon conduction band minima projections on the interface BZ overlap partially, to essentially the same extent for both Si(100) and Si(111) faces, as shown in Fig. 17. Thus the ballistic electrons, even when injected with a small transverse momentum component, will, if passing through a sufficient thickness of Au, reach the interface with substantial transverse momentum. It was concluded that the Au band structure itself could provide the required transverse momentum to facilitate interfacial ballistic transport in the Au/Si(111) system. Fig. 18 compares the theoretical calculations of Ref. [95] and the experimental data of Refs. [86,96] for the ratio of current injected in Au/Si(100) and Au/Si(111). Despite general agreement in the magnitude and in the overall trend with energy, a discrepancy up to 30% was obtained between the theory and the experiment. In some contradiction to the dependencies shown in Fig. 18, Weilmeier et al. [93] have found experimentally that BEEM current for Si(111) is more than 50% higher than that for Si(100) substrates for Au(111)/Si(111) and Au(111)/Si(100) studied at UHV conditions. To explain their findings, Weilmeier et al. suggested that in their experiment the Au(111) grains are aligned not in-plane but randomly with the Si(100) substrate.

Ripard et al. [97] have lately demonstrated the filtering effect of the Co/Cu interface by conducting a BEEM study on Co/Cu/Co trilayers. The strong reduction of the BEEM current when $\sim$2–3 Å of Co is evaporated on top of the Cu/Au/Si system was attributed to the band structure mismatch at the Co/Cu interface [Cu(111) has no propagating momentum states in a much larger cone around the film normal than Co(111)].

Recently, two cases of electron scattering (reflective vs. diffusive) in the metal have been considered by Menegozzi et al. [53] for the Au/Si system. A satisfying agreement of the model with

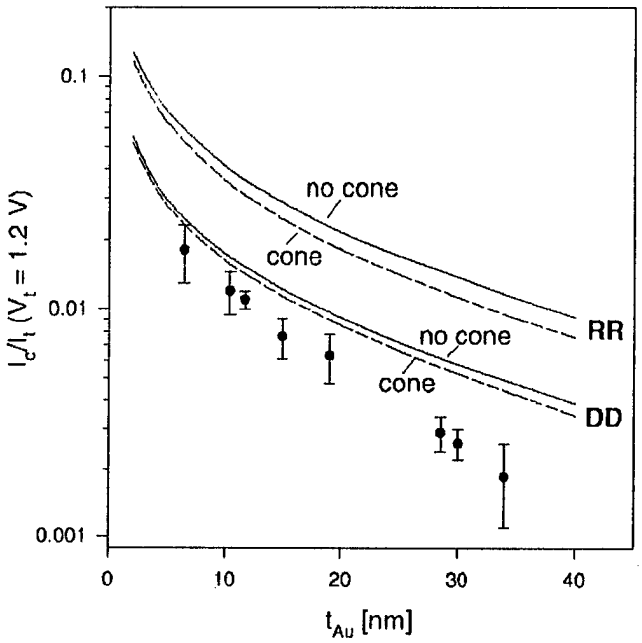


Fig. 19. The value of the BEEM characteristics I_c/I_t at the tunneling voltage of 1.2 V versus the Au film thickness t_{Au} for Au/Si(0 0 1). Curves are calculated for reflecting (RR) or diffuse (DD) interface structure as indicated by assuming an isotropical Fermi surface ("no cone", *solid line*) and a forbidden cone in the Fermi surface ("cone", *dashed line*). Measured data points are taken from Ref. [55] (from Ref. [53]).

the experiment on Au(111)/Si(111) was obtained, although it seems that a different interface scattering should be assigned to different experiments. The results for Au(111)/Si(100) exceeded at least by factor of 2 the experimental data. To improve the matching, the authors considered the effect of the Au band structure. However, the implementation of the Au(111) band structure gave only minor suppression of the BEEM current (~5%), as shown in Fig. 19, and thus cannot explain the mismatch.

Summarizing the results presented in this section, the effect of the hot electron scattering at the m–s interface was intensively studied both theoretically and experimentally. Experimentally, the BEEM transport was found to deviate significantly from the simplified picture of transverse momentum conservation at the m–s interface in the BK model, for the non-epitaxial Au/Si and Au/GaAs interfaces, and, to some extent, even for the more perfect epitaxial $CoSi_2$/Si system. Although certain upgrading in the theoretical description was achieved by elaborating the metal overlayer effect (e.g. metal band structure effect as well as effects of the electron scattering and multiple electron reflection in the metal layer on the BEEM transport), the proposed models suffered from the lack of the quantitative description of the experimental BEEM data. A real improvement of the theoretical fit to the experimental data was obtained by assuming a strong electron scattering at the m–s interface. Among the m–s interface scattering models, the MSIS model is the most developed model and allows a qualitative analysis of the electron scattering at the m–s interface. Using the MSIS model, the electron scattering probability was estimated to be ~85–95% at the non-epitaxial Au/GaAs interface.

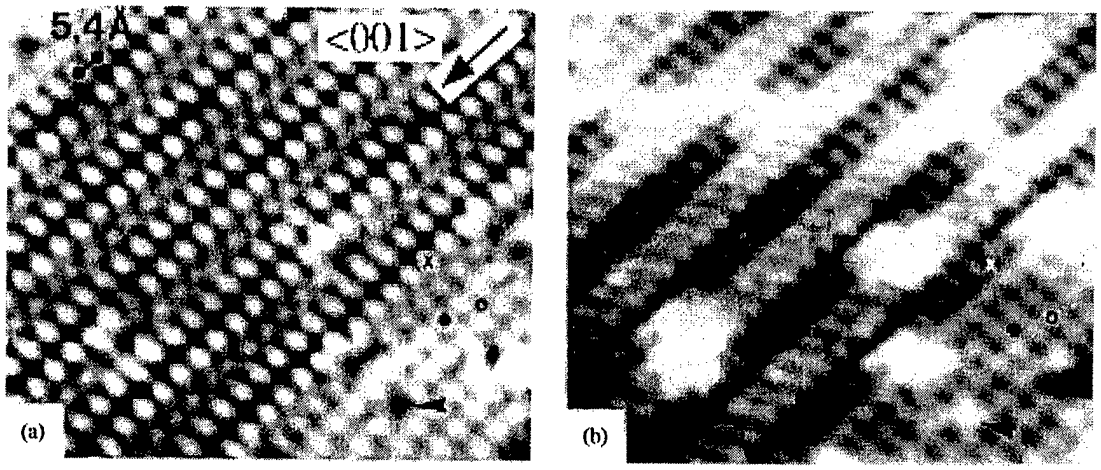


Fig. 20. STM topographic image (a) and simultaneously recorded forward BEEM image (b) on the Si-rich $CoSi_2$/n-Si (1 0 0) surface ($V_t = 1.5$ V, $I_t = 3$ nA, film thickness $d = 38$ Å). The center part displays a $3\sqrt{2} \times \sqrt{2}$ $R45°$ reconstruction, whereas the lower right is $\sqrt{2} \times \sqrt{2}$ $R45°$ reconstructed. The $\sqrt{2}$ topographic corrugation is 0.15 Å. The BEEM contrast ranges from 25 pA (black) to 55 pA (white) [98].

3. BEEM resolution

3.1. Spatial resolution

Sirringhaus et al. [8,98] have succeeded to image periodic surface structures at atomic resolution by in-situ BEEM on the Au/Si system, as shown in Fig. 20. The superior resolution (~10–20 Å) of the BEEM technique was demonstrated in imaging of microfabricated Au/SiGe diodes [99] and in imaging the Schottky barrier inhomogeneities for the $Au/Co/GaAs_{0.67}P_{0.33}$ system [100]. Also, a spatial resolution of ~10 Å (see Fig. 21) was obtained by in-situ BEEM in imaging the interfacial defects in 30 Å $CoSi_2$/Si(111) system at $V_t \sim 1.5$ V [77]. This resolution is reduced at high tip voltages due to the secondary electrons that have broad energy and momenta. For example, for $V_t = 6$ V, the spatial resolution of BEEM reduced to ~100 Å [101].

The observed high spatial resolution in BEEM experiments is to a certain extent an unexpected result. Indeed, the nanometer-scale resolution (~15 Å) was obtained after propagation through Au films as much as 100–150 Å in the case of the Au/Si system [102], whereas the free-electron model for the metals [56,102] predicts a BEEM resolution for relatively thick 100–150 Å Au films of at best ~100 Å. To explain the high spatial resolution observed in BEEM experiments, the metal band structure effect was examined as a possible candidate. In the case of epitaxial $CoSi_2$ layers on Si(111) substrates, Reuter et al. [9] showed by using the Keldysh Green-function method that hot electrons injected from a STM tip into a $CoSi_2$/Si(111) system form a highly focused beam due to the silicide band structure. It was calculated (see Fig. 22) that the current spatial distribution at $V_t = 1.5$ V remains highly focused (FWHM is 8.9 Å) after propagating through a 30 Å $CoSi_2$(111) film (the free-electron propagation in the metal would result in ~25 Å resolution). This spatial resolution is in a good agreement with the resolution of ~10 Å with which interfacial point defects

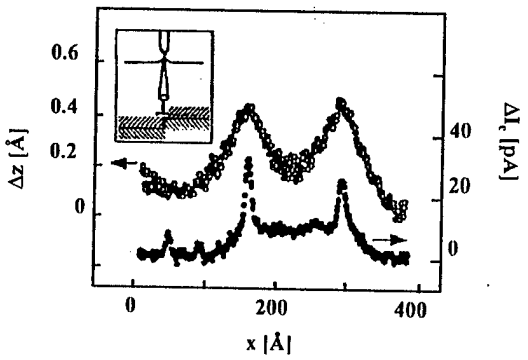


Fig. 21. STM topography (top) and BEEM (bottom) cross sections perpendicular to two parallel dislocation lines, together with a schematic drawing (inset) ($V_t = 1.5$ V, $I_t = 2$ nA, film thickness $d = 30$ Å) [77].

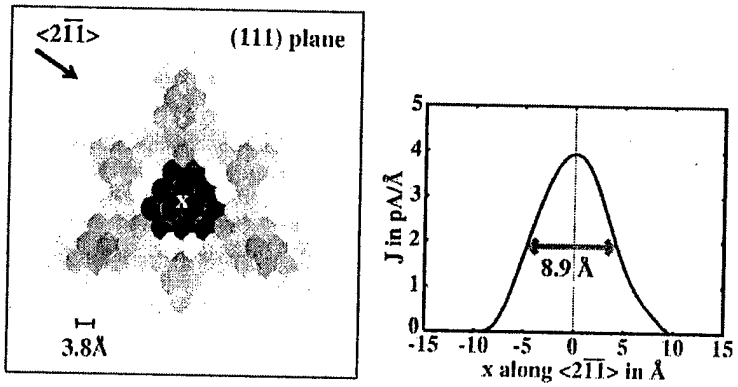


Fig. 22. Current distribution in a Si2 layer parallel to the surface after propagation through a 30 Å $CoSi_2$(111) film. Injection from the tip at 1.5 eV occurred in the center of the shown plane (white X), where the maximum current propagating in a focused beam along the ⟨111⟩ direction can still be found. The linear gray scale indicates current intensity at each atomic site: black maximum to white zero current. The right-hand panel displays a cut through the focused beam in the ⟨2-1-1⟩ direction from which a FWHM of 8.9 Å can be derived [9].

were resolved experimentally [77,81]. In the case of Au(111) on Si, the immediate consequence of the propagational gap of Au in the [111] direction is the opening up of the injected beam as the depth increases [95]. The deeper the chosen layer, the more spread over a larger ring (more precisely a spreading triangle considering the symmetry of a fcc (111) crystal) the current would be, but the sideward-directed propagation remains sharply focused (see Fig. 23). Thus, the real space distribution in any layer parallel to the surface would be a spreading triangle with the effective triangle side thickness of ~10 Å after propagating through ~60 Å of Au, in good agreement with the experiment [102]. Also, the results of the ensemble-Monte-Carlo simulations, where the effect of the metal band structure on the electron propagation was incorporated by cutting off the forbidden directions arising from gaps in the constant-energy surface, agree qualitatively with the (widely scattered) experimental BEEM data on the Au/Si system [103]. It should be noted though

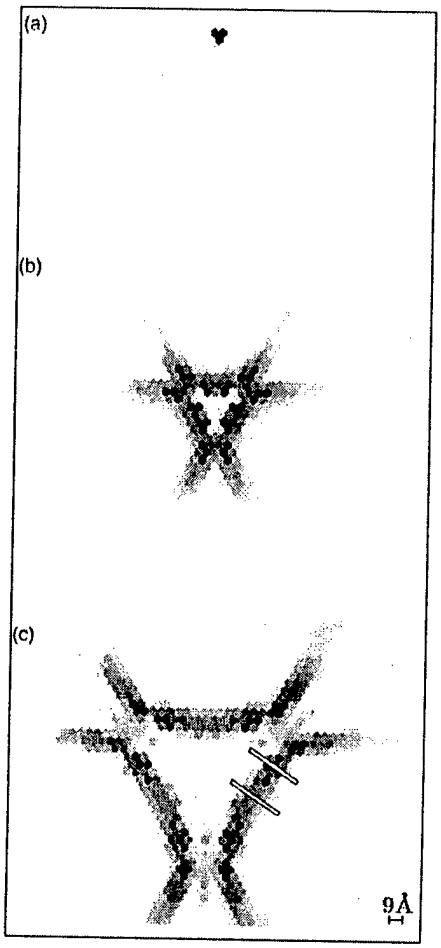


Fig. 23. Real-space BEEM current distributions for Au(1 1 1) after injection of current in one atom in the first layer (located in the center of the drawn layer). The tip is positioned at 5.0 Å height on top of the active atom, with $V_t = E_F + 1$ eV. Distribution in (a) 2nd layer (2.35 Å), (b) 10th layer (21.19 Å), and (c) 25th layer (56.51 Å). Each dot represents one atom in the corresponding layer and the gray scale indicates the amount of current passing through the atom: black for maximum current to white for zero current [95].

that, according to Fig. 23, the calculations predict "multiple images" of the interface objects, since a tip scan would sweep each of these lines across the interface object, each time leading to a focused signal. However, this effect was not observed experimentally.

3.2. Energetic resolution

3.2.1. Low-temperature BEEM

To fully exploit the BEEM capabilities, it is very beneficial to perform the BEEM experiments at cryogenic temperatures. There are at least three advantages to use the low-temperature BEEM technique: (i) improvement of the energy resolution due to the reduced thermal broadening of the

Fermi distribution, (ii) thermal noise is greatly reduced and (iii) thermal drift of the STM tip position is reduced.

The first successful BEEM operation at low temperatures (7 K) was reported in 1995 in the Au/Si(100) system [45]. It was shown that the vertical resolution was 0.3 Å and the positional drift of the microscope was less than 2 Å/h for operation at 7 K and 7 Å/h for operation at 77 K. In contrast, the usually observed room-temperature drift is much higher, of the order of several Å/min at $T = 300$ K [104].

Later, BEEM was performed on Au/GaAs [84] and Au/GaAs/$Al_{0.2}Ga_{0.8}As$/GaAs single-barrier structures [32] at $T = 7$ and 77 K. The SD-BEEM spectra (in the near-threshold region up to the onset of the L valley of GaAs) were used to analyze the electron-wave interference effect. The low background noise (<3 fA rms) and high spectral resolution ($k_B T \sim 0.5$ meV (6 meV) at $T = 7$ K (77 K)) enables the detection of features (additional peaks and dips) that cannot be reproduced by the Au/$Al_{0.2}Ga_{0.8}As$ single-interface model, since this model describes only the average shape of the SD-BEEM spectra. A more complete model [84] that incorporates the interference effects (reflections) at the Au/GaAs, GaAs/$Al_{0.2}Ga_{0.8}As$ and $Al_{0.2}Ga_{0.8}As$/GaAs interfaces is shown to describe the data more accurately and consistently over many spatial locations and samples.

As was pointed out by Henderson et al. [50], a main difficulty at very low temperatures (<10 K) is related to carrier freeze-out in the semiconductor, which leads to charging effects in the collector electrode. In contrast to bulk samples, satisfactory BEEM measurements in liquid He can be obtained with the use of a low-dimensional electron gas layer as a collector electrode. Eder et al. [46] have studied the laterally patterned quantum wires fabricated on modulation-doped GaAs/AlGaAs heterostructures by BEEM at $T = 4.2$ and 300 K, using a 2D electron layer at the AlGaAs/GaAs interface to collect the BEEM current. At $T = 4.2$ K, no decrease of STM resolution was found in comparison with room-temperature measurements.

3.2.2. *Semiconductor STM tips for BEEM*

The use of degenerate semiconductors as the STM tip is potentially very useful to increase significantly the BEEM energy resolution, since the electron distribution is limited to the small range between the conduction band edge and the Fermi level. This allows a quasi-monochromatic energy distribution of the injected electrons, in contrast to the case of the conventional metal tips where the energy distribution of injected electrons is defined by the combined effect of the tunneling probability and the wide Fermi-distribution function of the tip electrons. In the literature, degenerate semiconductor tips such as highly doped GaAs and Si were already used for STM experiments [105,106]. Recently, BEEM experiments with n-doped InAs tips (InAs is a very promising material for semiconductor tips since it has a natural surface accumulation layer of electrons [107]) were reported at room temperature [108,109]. The InAs tips were fabricated by cleaving the InAs wafer, and then the cleaved piece of InAs was glued on the tilted holder so that one corner of the InAs piece can be used as a tunneling tip. In the experiment, BEEM spectra of modulation-doped Au/GaAs/AlGaAs sample at 300 K and in liquid helium obtained with InAs tip are of comparable amplitude with the spectra obtained using standard Au tips. While the spectral shape of the BEEM spectra obtained with InAs tips is somewhat different from the reference curves measured with Au tips, the expected enhanced energy resolution was not observed with InAs tips. Possible reasons for this were attributed to hot electron scattering processes in the Au film. We

believe that future BEEM study of structures, which give qualitatively different BEEM spectral shapes for metal and semiconductor STM emitters (for example, buried resonant heterostructures) should provide an experimental support for the energy-resolution effect of semiconductor STM tips.

3.3. Depth resolution

The hot electron transport in BEEM experiments can be considered to be ballistic only with certain restrictions. The real picture is certainly different because of consecutive scattering processes in the metal base film, at the m–s interface and in the semiconductor itself. Even for the case of a perfectly clean device and extremely low temperatures, the carrier mean free path is usually comparable with the device thickness. Therefore, the ballistic description of the electron transport is restricted to layers close to the surface. As a consequence, BEEM can be characterized by its depth resolution.

Monte Carlo simulations [110] of electron transport in the Γ, L and X valleys of GaAs were performed using the MSIS model [85]. With increasing object depth beneath the Au/GaAs interface, the Monte Carlo simulations predict (a) significant cooling of hot electrons, on the order of ~ 3 meV/nm at 1.5 eV, and (b) significant redistribution of electrons among the conduction bands (see Fig. 24). In order to show experimentally that scattering inside the semiconductor structure affects the BEEM spectra, in Ref. [5] we compared the BEEM spectra for several pairs of Au/GaAs/AlGaAs single barrier (SB) samples with the same SB composition but with different cap thickness of 50 and 300 Å. The obtained SD-BEEM spectra of the $GaAs/Al_{0.2}Ga_{0.8}As/GaAs$ SB samples are shown in Fig. 25 for $T = 85$ and 300 K (these data are representative for all samples pairs). Since the expected mfp lengths for the Γ, L and X electrons at $T = 300$ K are ~ 1000, ~ 100 and ~ 10 Å, respectively [43], manipulations with the heterostructure thickness in the range of 100–300 Å should affect mainly the L-electrons contribution to the BEEM current (the effective heterostructure thickness, that will affect the BEEM current collection, is the combined thickness of the GaAs cap layer and the SB layer). Indeed, as one can see from Fig. 25, the L-electrons contribution for the sample with 300 Å-cap layer (with the SB thickness of 50 Å, the total effective heterostructure thickness is ~ 350 Å) is reduced by factor of 3 as compared with the 50 Å-cap layer sample (the total effective thickness of ~ 100 Å), whereas the Γ contribution has remained intact. As the temperature decreases from 300 to 85 K, in addition to the spectrum shift expected from the temperature dependence of the energy gap, a strong increase of the signal is observed for the L-electrons in the SB sample with the 300 Å-cap layer, whereas the SD-BEEM spectral shape remains essentially the same for the SB sample with the 50 Å-cap layer. These experimental results are in agreement with the expected increase in the mfp of the electrons with the decreasing temperature. As the temperature decreases from 300 to 85 K, the calculated mfp near the energy threshold increases from ~ 1000 to ~ 1500 Å for Γ electrons and from ~ 100 to ~ 300 Å for the L electrons [43]. The X-electrons contribution is completely obscured in the room-temperature BEEM spectra presented in Fig. 25 due to their effective scattering in the GaAs cap layer. Indeed, the X-electrons mfp (~ 10 Å at $T = 300$ K and ~ 30 Å at $T = 85$ K) is shorter than the effective heterostructure thickness for all samples (≥ 100 Å). Thus, while the initial electron distribution among the conduction bands of the semiconductor is specified by transport through the metal base layer and by the m–s interface scattering, further electron transport is governed by the difference in

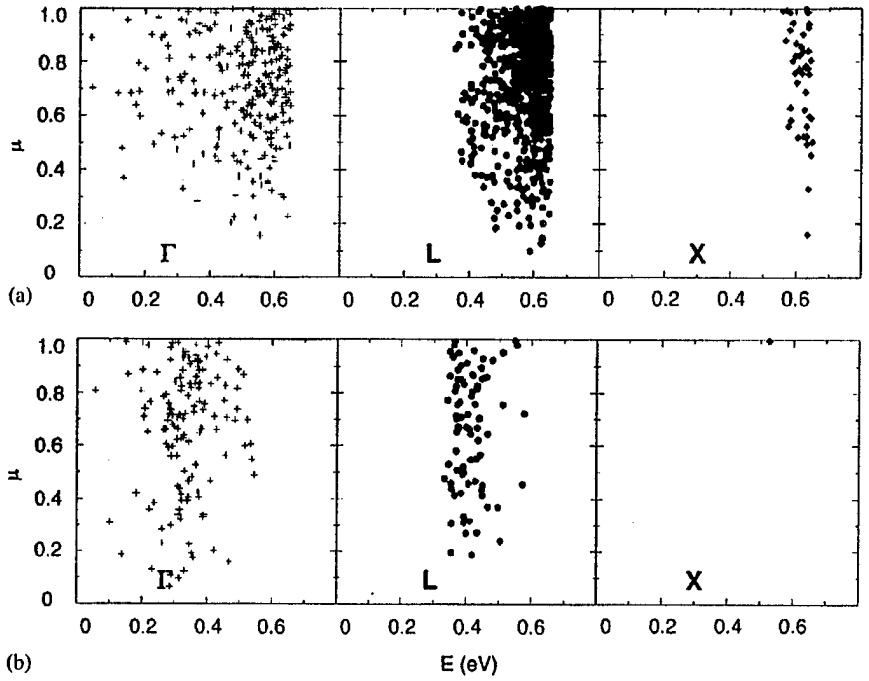


Fig. 24. For the populations of the electrons in GaAs resulting from a BEEM experiment of Au–GaAs, the energy and angular distributions are shown. The y-axis is the cosine, denoted by μ, of the angles of the velocities of the electrons with respect to the interface normal and the x-axis is the kinetic energy E of the electrons with respect to the Γ-valley minimum. The condition is the strong scattering limit of the MSIS theory at 300 K and at a bias of 1.53 V. Initially, 1000 electrons are injected at the Au–GaAs interface. (a) The initial electron distributions of Γ-(left panel and crosses), L-(middle panel and circles), and X-(right panel and diamonds) valley electrons at the Au–GaAs interface. (b) The distributions at 600 Å below the Au–GaAs interface (from Ref. [110]).

the electron mfp length for the Γ, L and X electrons. In imaging the buried semiconductor structures, the mfp length (in the case of GaAs-based structures, the mfp length of Γ-electrons as the longest one) will determine the depth resolution of BEEM. It implies that the BEEM contrast will be not enough to resolve a semiconductor structure buried much deeper than the electron mfp length ($\sim$1000 Å for the Γ-electrons of GaAs). Therefore, as the heterostructure, characterized by several transport channels, is buried deeper, the information about these conduction channels is gradually gone, starting from the transport channel with the shortest electron mfp length.

Lee et al. [111] have performed Monte Carlo simulations of the transport of electrons injected into the Γ valley of GaAs for BEEM imaging and spatially resolved spectroscopy of model quantum dots (QD) and quantum wires (QW) buried beneath the Au/GaAs interface. For Monte Carlo simulations, a modified BK model was used, with the buried object forming a local barrier for the electron transmission (such as AlSb and GaSb dots grown on GaAs). The collection plane was located at 2000 Å below the Au/GaAs interface, and the QD (QW) was modeled as a thin disk 300 Å in diameter (a thin 300 Å wide stripe), which specularly scatters electrons of all energies. Fig. 26 shows the calculated BEEM current as a function of lateral displacement between the STM tip and the center of the buried dot (wire), for three depths of the QD (QW) from the Au/GaAs

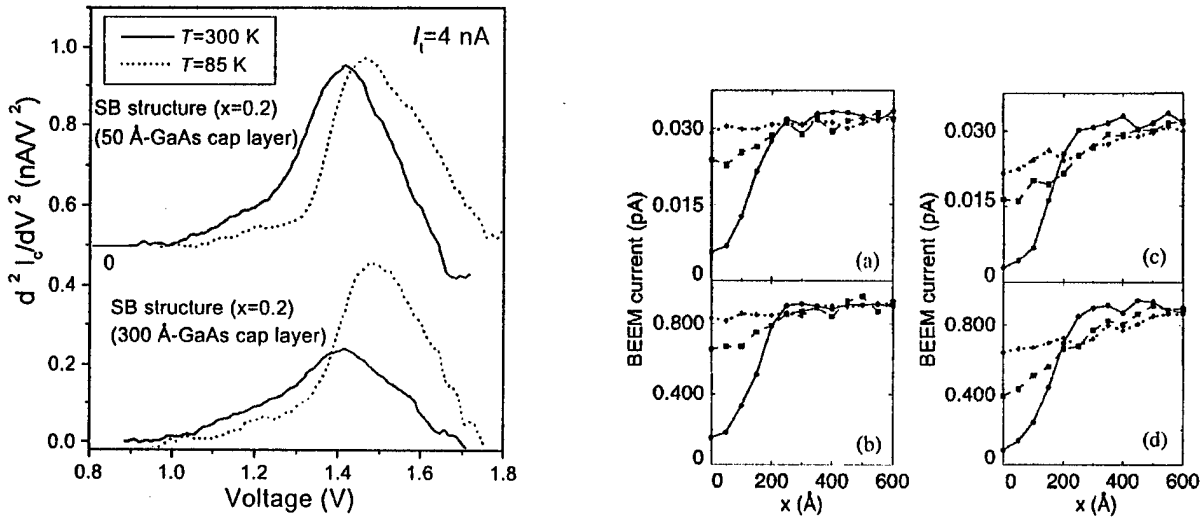


Fig. 25. SD-BEEM spectra for $GaAs/Al_{0.2}Ga_{0.8}As$/ GaAs single barrier (SB) samples with 50 and 300 Å GaAs cap layers, taken at $T = 300$ K (solid curves) and $T = 85$ K (dotted curves). For clarity, the spectra are shifted along the vertical axis [5].

Fig. 26. The BEEM current as a function of lateral displacement at a tip-to-sample bias of 1 V (a) and 1.25 V (b) for a perfectly reflecting quantum dot 300 Å in diameter, and at a tip-to-sample bias of 1 V (c) and 1.25 V (d) for a perfectly reflecting 300 Å wide quantum wire. The lines are drawn as a guide for the eye. In all cases, the solid line (circle) is for $d = 100$ Å, the dashed line (square) is for $d = 300$ Å, and the dotted line (diamond) is for $d = 600$ Å [111].

interface. One can see that the BEEM contrast decreases with the depth of the object. At a depth of 600 Å, a QD is hardly detectable by BEEM whereas the QW still shows significant contrast, due to the added dimension. In addition, it was shown that due to an interplay between a geometric filtering effect and the role of the electron–phonon scattering (assuming electron transverse momentum conservation at the m–s interface), there is a range of optimum depths for the sharpest crystal momentum and energy distribution of the electrons incident upon the buried structure.

As an example of BEEM imaging of deeply buried localized objects, a study of misfit dislocations at the $In_xGa_{1-x}As/GaAs$ interface buried 400–700 Å below the m–s interface was conducted in our laboratory [42,43,112]. The cross-hatch misfit dislocation network in $In_xGa_{1-x}As/GaAs$ was ideally suited for BEEM as the dislocation core beneath the surface gives rise to a perturbation at the surface easily visible by AFM or STM. Samples consist of $In_xGa_{1-x}As$ of different thickness grown on (001) GaAs substrates and were designed such that the dislocations will exist (Indium concentrations of $x = 0.15$–0.3 and InGaAs thickness of $d = 400$–750 Å were examined [112]). Simultaneous STM and BEEM imaging, shown in Fig. 27, revealed a decrease in the BEEM current directly associated with the cross-hatch pattern in the STM. The minimum in the BEEM current, through analysis of line scans, was found to be shifted $\sim$400 Å from the maximum of the STM image, indicating that the loss in BEEM current is not simply a surface affect, but one that arises from the dislocation core located $\sim$600 Å beneath the surface (see Fig. 28). This is in agreement with the theory that the dislocation core comes to the surface not along the growth

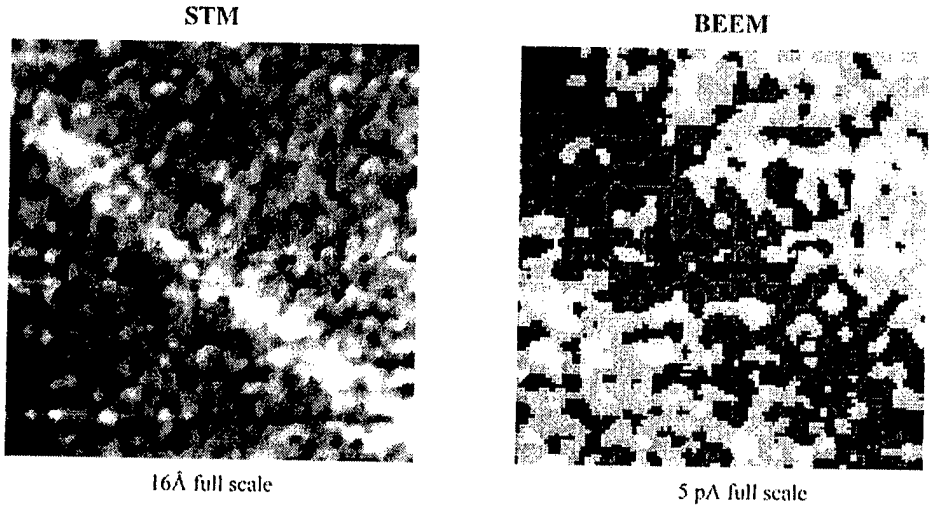


Fig. 27. 1 μm × 1 μm STM (left) and BEEM (right) images of $In_xGa_{1-x}As/GaAs$ heterostructure ($x = 0.2$, $d = 600$ Å). The STM image shows the growth perturbation caused by the dislocation core. The BEEM image shows the decrease in current at the presence of the cross-hatch [112].

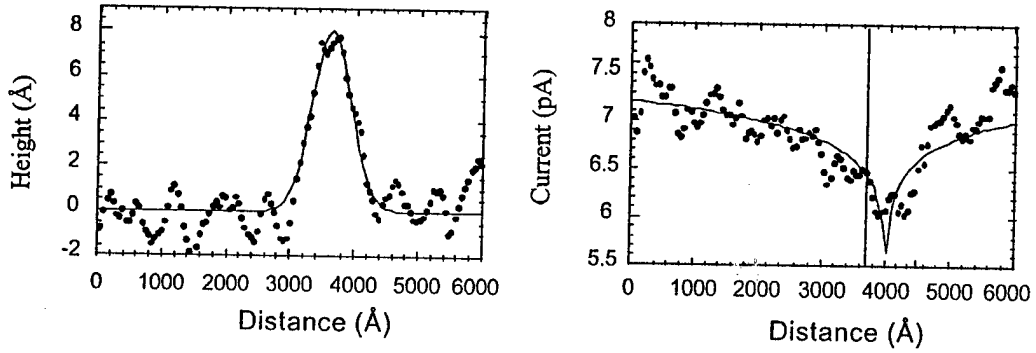


Fig. 28. STM (left) and BEEM (right) linescanes taken over dislocation images. The STM linescanes show a Gaussian peak while the BEEM linescanes show a drop in the BEEM current shifted from the peak in the STM image [112].

direction, but along the glide plane (111), which is 54.7° off of the (100) direction. Monte Carlo calculations were used to understand the mechanism responsible for this decrease [112]. The simulation of the charged model fit closest to the experimental data suggesting that backscattering off of line charges to the metal base is responsible for the loss in BEEM current.

4. Homogeneity of BEEM transport

4.1. BEEM dependence on the m–s interface quality

The importance of interface chemistry for BEEM was underlined by Fernandez et al. [4], who showed on example of Au/Si interfaces that when Au is deposited in-situ on a UHV-cleaned Si

surface, no BEEM current can be detected through the resulting Au/Si interface. This was explained in terms of additional scattering from the Au–Si disordered alloy formed by Au–Si interdiffusion. The interface transmittance can be increased to some extent by HF-treatment of the substrate surface prior to the Au deposition. In this case, the residual impurity at the interface can act as a barrier, which passivates the Au–Si interaction enhancing the ballistic transmittance. For the Au/GaAs(100) system, Hecht et al. [113] showed that the MBE termination of the GaAs surface by a 2 monolayer (ML) AlAs diffusion barrier prior to the Au deposition results in a significant increase of the interface areas, which support BEEM transport, and produces a high-quality Schottky barrier.

It is well known that the Schottky barrier height depends on the chemical and mechanical pretreatment during the diode preparation. Chemical pretreatment in aqueous HF or HCl solutions and subsequent rinses in deionized water resulted in the Schottky barrier height variations up to 10–20% in Au/III–V (GaAs, AlGaAs, InP and InGaAs) [114] and Au/Si [115]. Everaert et al. [116] have found by using BEEM that mechanical polishing of the etched GaAs substrate prior to the Au deposition produces an increase of the Au/GaAs Schottky barrier. This increase was explained by the passivation of the interfacial defects by hydrogen and the consequent pinning of the Fermi level. Schottky barrier variations in Au/n-GaAs due to reactive ion etching by using $SiCl_4$ was studied by Vanalme et al. [117] who found the Schottky barrier reduction (by $\sim 6\%$) due to the change of the stoichiometry of the surface region produced by the reactive ion etching. In addition, Stockman et al. [118] showed that the BEEM current homogeneity and stability over time is very sensitive to the interface flatness, namely that the interface formed on the flat (110) cleavage plane of a GaAs(100) wafer, is more homogeneous than the interface formed on the as-grown (100) plane and transmits ballistic electrons for months without any significant deterioration. In general, the Schottky barrier homogeneity is a very important device characteristic, since in its absence, the integrated macroscopic Schottky barrier height extracted from the I–V measurements might be quite different from the averaged BEEM threshold [119].

4.2. *BEEM dependence on the surface topography*

In the case of epitaxial $CoSi_2$ layers on Si, the average BEEM current depends on the surface reconstruction [120]. On the atomic scale, the BEEM corrugation was out of phase with the STM topography. This surface-induced contrast of BEEM was attributed to the atomic-scale variations of the energy distribution of tunneling electrons (local variations of the tunneling barrier). The in-situ BEEM measurements [93] of Au/Si(111) and Au/Si(100) interfaces show the BEEM current variations that correlate with the surface slope (larger BEEM current at atomic steps and grain boundaries), as shown in Fig. 29. This result was explained in terms of the larger density of surface states at flat Au surfaces. Electrons, tunneling from the STM tip to such surface states, will have a substantially lower probability of reaching the interface before scattering inelastically than those electrons that tunnel directly to empty bulk states. It was noted that the BEEM contrast effect has generally not been seen previously for Au on Si, possibly due to the morphology or orientation of the Au grains in those studies [121].

Recently, Weilmeier et al. [122] have successfully imaged the variations in the density of unoccupied surface states by conducting BEEM on Cu/Au, Ag/Au and Au/Cu/Au (111) overlayers UHV-grown on Si(100) and Si(111) substrates (a second (Cu or Ag) film overlayer, typically

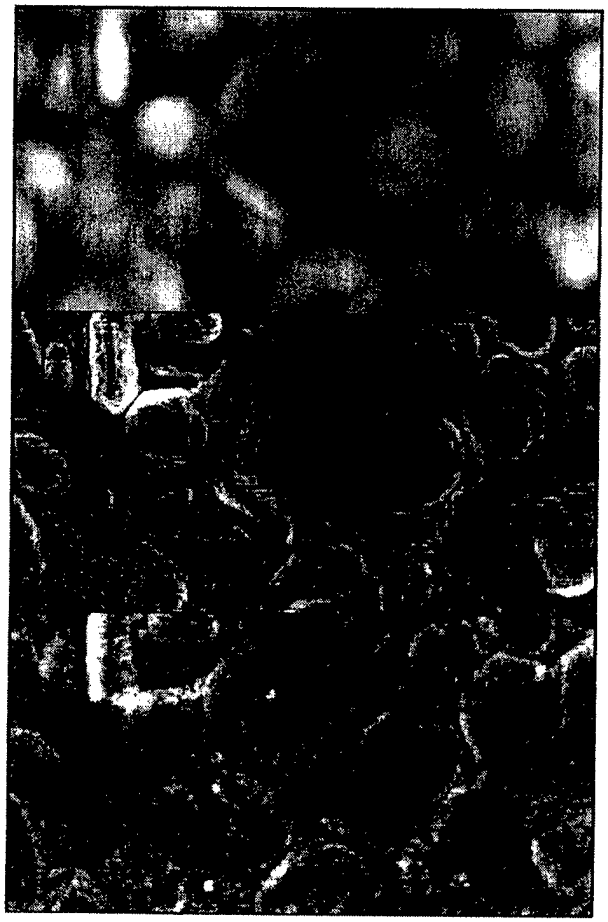


Fig. 29. STM and BEEM images of a 200 Å Au/Si(100) sample. The 100 nm × 350 nm images are obtained at $V_t = 1.1$ V and $I_t = 1$ nA. The top image is the STM surface height (17 Å white to black), the middle image is its calculated surface slope, and the bottom image is a map of BEEM current, I_c. The mean BEEM current is 2.25 pA, and total range shown is 2.4 pA. BEEM contrast is about 30% [93].

30–100 Å thick, was deposited on an initially deposited 100 Å Au(111) layer). The BEEM results were confirmed by STM spectroscopy that yields a high density of the conduction-electron surface states on the flat areas of the spiral-like Cu/Au,Ag/Au and Au/Cu/Au (111) grains. The BEEM current increased on grain boundaries where the density of the surface states is small. This effect was explained to be due to the preferentially (111) oriented growth of the Au film under UHV conditions. As a consequence of the Au band structure restrictions on the electron propagation in the normal-to-surface direction, higher current flows through the unoccupied surface states.

With polycrystalline metal films on GaP(110) [7,123] and GaAs(110) [118] substrates, the BEEM current was attenuated at grain boundaries, and this effect was attributed to elastic scattering decreasing the transmission rate into the semiconductor conduction-band-centered substrate. In those cases the metal band structure did not forbid electron transport normal to the surface, and thus there was no involvement of surface states.

4.3. BEEM-induced modification of carrier transport

The variation of interface transmission properties under hot electron injection was reported in BEEM experiments. Fernandez and Hellen et al. [4,124] first reported that the transmission properties of some specific Au/Si interfaces could be modified on the nanometer-scale when the energy of injected electrons was a few eV higher than the Fermi level of the metal base. It was pointed out that irreversible modification in the transmission rates for the buried interface can occur when the system is stressed with electrons injected at $V_t > 3$ V above the Schottky barrier. Spatially, the modification typically consists of a region of decreased transmittance of a few

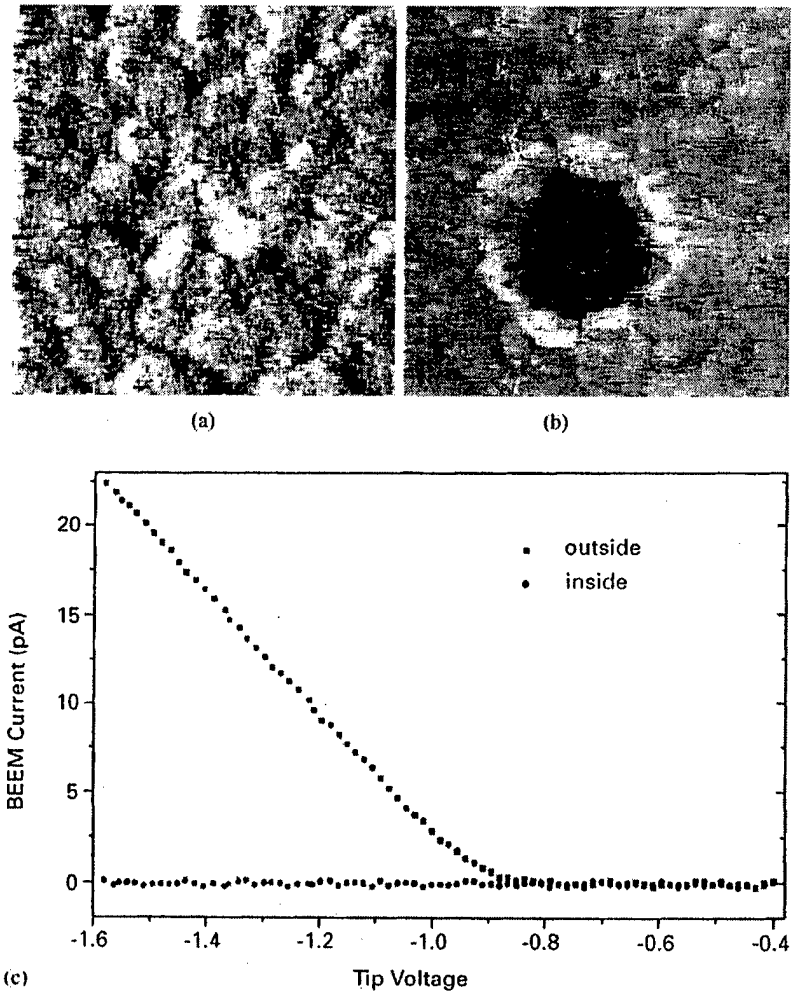


Fig. 30. Topographic and simultaneously acquired BEEM images with a tip bias $V_t = -1.30$ V and tunneling current $I_t = 2.5$ nA of an Au/n-Si(1 1 1) sample after modification operation. The image area is 800 Å × 800 Å. The applied voltage pulse is −4.0 V with a duration of 10 s. (a) STM topography of the Au surface. The gray-scale range is 20 Å. (b) Corresponding BEEM current image. The gray scale covers a range of 3–30 pA. (c) Spectroscopic measurements performed inside and outside the modified region reveal significant differences [125].

hundred Å in diameter surrounded by a ring of increased transmittance. A model was proposed in which the tunneling current at high bias enhances diffusion at the interface resulting in a change in the local atomic distribution.

Qiu et al. [125] reported on the observation of a series of fabricated interfacial features with significantly different characteristics. Under applied negative tip bias $V_t > 2.5$ V for several seconds in the Au/n-Si(100) and Au/n-Si(111) systems, a strong modification of the BEEM image occurred without noticeable change in the surface topography, as shown in Fig. 30. At the center of the modified zone there is a dark "hole" with zero BEEM current surrounded by a ring with enhanced BEEM current. The area of this hole increased nearly linearly with the exposure time (diameter increased from ~300 Å to ~450 Å when exposure time increased from 5 to 30 s at $V_t = -4$ V). A threshold pulse voltage of about 2.5–2.8 V has been observed in such a modification. When the electron energy was lower than this value, little change could be observed in the interface BEEM image, even with $I_t = 12$ nA and exposure time of several minutes. Thus the mechanism responsible for the modification requires an activation energy with the following accumulation of injected electrons (this is evident from the increase of the modified area with the exposure time). Also, surprisingly, when a voltage pulse of higher amplitude, $V_t = -5$ V for 20 s, was applied, the modified area shows higher transmittivity than the unmodified surface (Schottky barrier remains essentially the same). The authors concluded that the mechanisms proposed earlier (stacked Au terraces and Au/Si interdiffusion [124]) to explain the interface modification are not sufficient to account for all the observations.

Recently, Hasegawa et al. [121] have conducted BEEM study on Au/Si(111) where the Si substrate was chemically treated by dipping into 5% HF solution prior to gold deposition. They have shown that by applying a negative voltage of $V_t > 3$ V on the tip for several seconds, a local region was created where no BEEEM current flows at the interface. The modified area can be erased by applying a voltage with the opposite polarity (see Fig. 31). It is found that the minimum size of writing and erasing corresponds to Au grains, suggesting a method of rewritable memory on a nanometer scale. The possible origin of this effect is due to vacancies or due to the charging effects at the interface.

5. BEEM on buried semiconductor heterostructures

Although originally the BEEM technique was invented as a unique microscopic and spectroscopic method to probe the Schottky barriers on a local scale, the BEEM technique can be successfully used to study the electronic properties of *buried* heterojunctions. If there is an additional potential barrier (higher than the Schottky barrier) in the path of the injected carriers, the threshold should correspond to the height of this barrier. This is the essence of the heterojunction experiments using BEEM to probe the semiconductor band structure (spatially) beneath the m–s interface. The modeling of the BEEM transport in interference filter structures clearly demonstrated the potential of BEEM to probe buried quantum structures [126]. Recent research has proven that BEEM is a very powerful method for imaging and electron spectroscopy of buried defects and low-dimensional quantum structures.

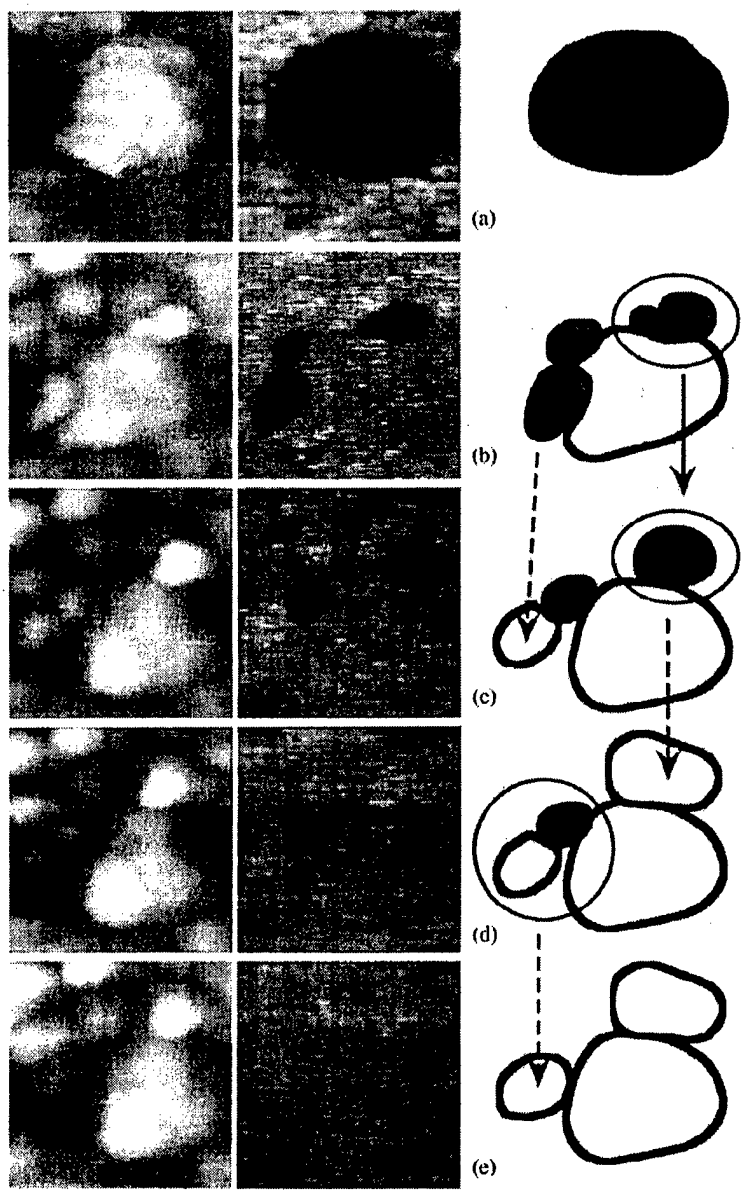


Fig. 31. Series of STM (left) and BEEM (middle) images during an erasing process of the Au/Si(1 1 1) interface. (a) Images taken before the process and (b)–(e) images taken after every scanning with a positive voltage of 1.34 V. The size of the images is 47 nm × 47 nm. Schematics shown on the right demonstrate a variation of Au grain structures with the BEEM current contrast together [121].

5.1. BEEM in simple metal/semiconductor structures

5.1.1. Metal/II–VI- and metal/IV-semiconductor systems

So far, only few BEEM studies were reported on metal/II–VI semiconductor systems. The BEEM technique was applied to measure the Schottky barrier contacts fabricated by evaporating

Au on as-grown n-ZnSe [104] and on chemically etched n-ZnSe [127] and CdTe [119]. BEEM studies in as-grown-prepared Au/ZnSe give relatively narrow Schottky barrier height distribution (1.32–1.43 eV) [128], in contrast to wide Schottky barrier height distribution (1.53–2.15 eV) found in Au/ZnSe diodes prepared by chemical etching [129].

For metal/IV semiconductor systems, one of the first applications of BEEM was the nanometer-scale study of ballistic transport in the metal/Si Schottky barrier system [23], and today this system is probably the most studied system in BEEM experiments and model calculations. Some BEEM results for this system were reviewed in Sections 2 and 3 of the current paper (see also previous review articles [2,3]).

Recently, BEEM study of group-IV semiconductors was extended to SiGe and SiC compounds. Bell et al. [12,130] had applied BEEM to study the effect of strain in metal/$Si_{1-x}Ge_x$/Si structures. The 500 Å $Si_{1-x}Ge_x$ layers ($x = 0.18$ and 0.25) were pseudomorphically strained (well below the critical thickness). The presence of strain modifies the $Si_{1-x}Ge_x$ band structure. Whereas Au/ Si(100) BEEM spectra show a single threshold, the Au/$Si_{1-x}Ge_x$/Si BEEM spectra usually exhibited two thresholds (a splitting of $\sim$0.3 eV at $x = 0.25$). These thresholds were attributed to the strain-induced conduction band splitting of $Si_{1-x}Ge_x$. The threshold positions as a function of the splitting were found to follow the theoretical work of Van de Walle and Martin [131], who predicted that the four strain-equivalent conduction band minima in the layer plane are lowered in energy, while the two minima normal to the plane are raised in energy. Based on BEEM results, the spatial distribution of strain in $Si_{1-x}Ge_x$ layers was found to be very heterogeneous (conduction band splitting varied from 0 to 0.35 eV from point to point across the sample surface). This characteristic was explained in terms of the Au/$Si_{1-x}Ge_x$ interface roughening by Au–Si interdiffusion. Indeed, in the case of Ag/$Si_{1-x}Ge_x$ where the interface is smooth, the observed splitting was uniform [130].

In the case of SiC, which is a technologically important IV–IV compound wide-bandgap semiconductor, the Schottky barrier height was studied for (Pd,Pt)/6H-SiC (1.27 eV, 1.34 eV) and (Pd,Pt)/4H-SiC (1.54 eV, 1.58 eV) [21,22]. In 4H-SiC, an additional CBM was observed at $\sim$0.14 eV above the lowest conduction band minimum. Enhanced ballistic transmittance (without significant change of the Schottky barrier) was observed over a region intentionally stressed by injecting high-kinetic-energy hot electrons (10 eV above the Fermi level) using BEEM in the Pt/4H-SiC system.

5.1.2. Metal/III–V semiconductor systems

The BEEM technique was extensively applied to study Schottky barriers and hot electron transport peculiarities in many metal/III–V semiconductor structures, e.g. Au/GaAs [5, 29, 49, 84, 116], (Au,Mg,Cr)/GaP [7,64], Au/InAs [132–134], (Au,Co)/$GaAs_{0.67}P_{0.33}$ [100] and Al/AlSb [135]. In this section, we will discuss the very recent BEEM results in several important compound semiconductors such as GaN, GaInP and GaAsN.

5.1.2.1. GaN. Although GaN exhibits high-efficiency optical performance despite its extremely high dislocation densities [136], the dislocations are very likely to have significant electrical and structural impact on high-power electronic devices.

Recent study of Brazel et al. [18] has revealed high BEEM current densities and low effective Schottky barrier heights in small localized areas around dislocations with a screw component. The GaN-films used in that study were grown by metal organic vapor deposition (MOCVD) at

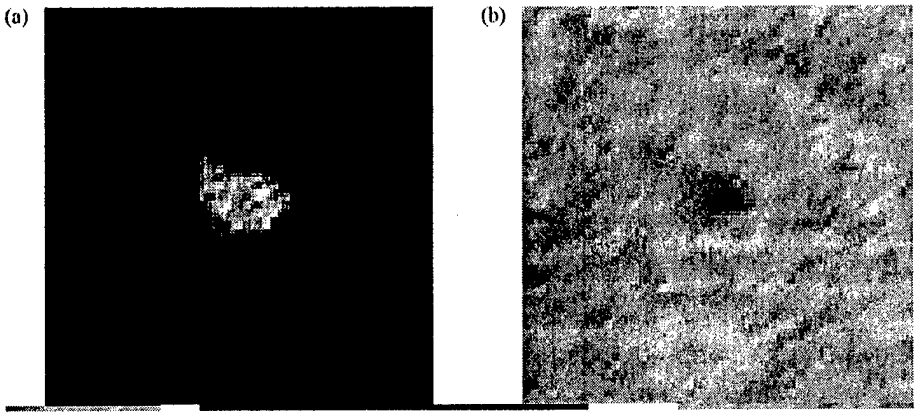


Fig. 32. Collector current image (a) and simultaneously taken STM image (b) of the 70 Å Au cap layer of a 2500 Å × 2500 Å area ($I_t = 1$ nA, $V_t = -1.8$ V). The small high-current area (~ 500 pA, bright spot) surrounded by areas with zero collector current (< 0.2 pA) shown in (a) is typical for the investigated GaN-films. The Au layer (right) exhibits a pit (dark) at the location of the high-current area (bright spot (a)) [18].

atmospheric pressure. For these measurements unintentionally doped, n-type GaN-films of thickness ranging from 1 up to 3 μm were deposited on *c*-plane sapphire substrates. Fig. 32(a) shows a typical collector current distribution for the MOCVD-grown GaN films deposited on *c*-plane sapphire substrates. From most parts of the sample no collector current was detectable (dark area), and only in small localized spots, like the shown light area, high collector currents could be observed. In this representative case, the measured current was ~500 pA, which is half of the injected 1 nA tunneling current. To localize the high current areas on the sample surface and possibly correlate them with surface features, images of the current distribution and the topography were taken simultaneously. Fig. 32(b) shows the STM image of the Au cap-layer, which corresponds to the collector current image in Fig. 32(a). The 70 Å thick Au-layer exhibits a ~500 Å wide pit at exactly the same position of the high current area in the collector current image. An AFM-image taken of the same sample shows several of these pits that represents the surface terminations of mixed and pure screw dislocations. Thus, it was concluded that the high BEEM current areas are located at dislocations with screw character. To get information about the Schottky barrier heights, the BEEM spectroscopy in both high and low currents areas was carried out. The low current areas do not show any measurable collector current, whereas for the high current areas, we have found barrier heights in the range from ~0.3 up to ~0.95 eV. The threshold values vary within one high-current area and from area to area. In addition, it was possible to reverse the tunneling voltage and inject holes into the very same high current areas as well, with the observed threshold energy of ~0.3 to 1.3 eV, as shown in Fig. 33. The observed thresholds for carrier injection do not seem to be typical Schottky barriers, but rather capturing barriers of trap states in the vicinity of the dislocations. These results definitely suggest that acceptor- and donor-like trap states coexist in the same area, and are being filled during carrier injection into the GaN. The maximum threshold energies observed for electrons (0.95 eV) and holes (1.3 eV) add up to 2.25 eV, remarkably close to the parasitic yellow luminescence commonly observed in GaN. In the future, we plan to conduct more experiments to answer the question whether the same

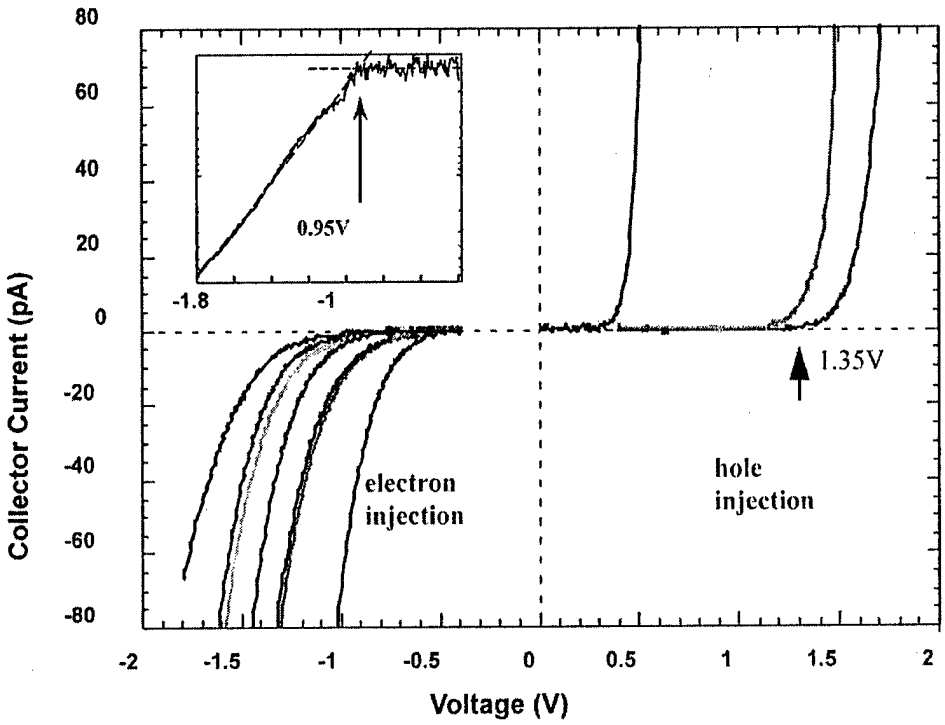


Fig. 33. BEEM spectra for electron and hole injection. Log-plots like the one shown in the inset were used to approximately determine the threshold voltage necessary for current injection into the high-current areas [18].

electronic states, which are causing the high current densities are also responsible for the yellow luminescence.

Bell et al. [17,19,137] also studied metal/GaN Schottky barriers. The GaN layer of ~ 2 μm thickness was grown by MOCVD on a (000 1) sapphire substrate. After several attempts, they succeeded in measuring the BEEM current in the Au/GaN and Pd/GaN systems if the GaN surface had been spin etched using 1:10 HCl:ethanol solution prior the metal deposition. Two thresholds were observed in the BEEM spectra (~ 1 and 1.2 eV) [17]. The first threshold was interpreted to be the Schottky barrier (consistent with the conventional macroscopic I–V measurements), while the origin of the second threshold might be due to the strain-induced conduction band splitting. Imaging of the Au/GaN interface reveals transmission in nearly all areas, although the magnitude is small and varies by an order of magnitude (0.2–2 pA at $I_t = 2$ nA), as shown in Fig. 34. It was shown that if the HCl/ethanol treatment of the GaN surface is followed by sample annealing in UHV or in N_2 prior Au deposition, much higher and more uniform transmission across the Au/GaN interface can be achieved [19]. However, while the GaN annealing resulted in more than one order of magnitude increase of the BEEM current, it also resulted in undesired alteration of the Schottky barrier height. Fig. 35 shows to reduction of the Schottky barrier from ~ 1.06 V to ~ 0.92 and to ~ 0.69 V after UHV annealing for 15 min at 340 and 580°C, respectively. These barrier height changes were interpreted in terms of creation of vacancies or their diffusion toward the GaN surface.

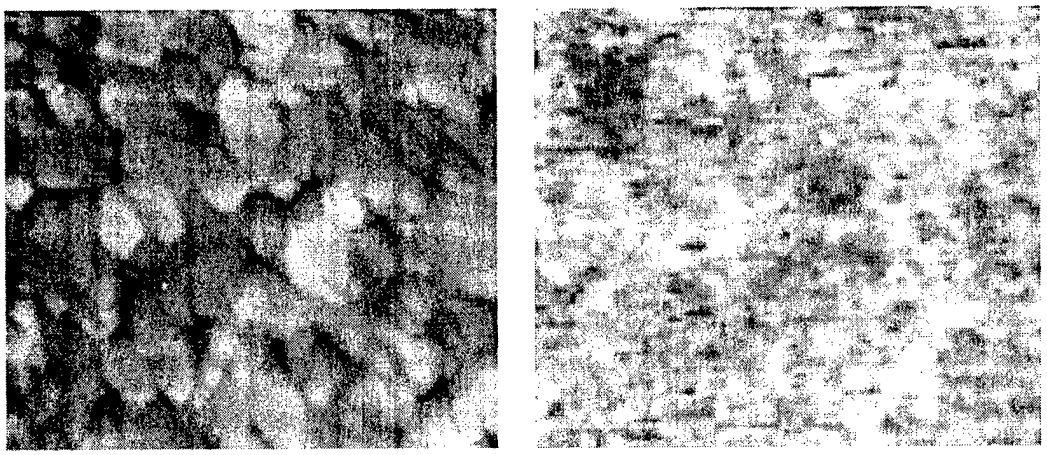


Fig. 34. STM topography/BEEM image pair for a Au/GaN sample. The topography of the Au surface was obtained at $V_t = 0.5$ V, $I_t = 1$ nA, and the BEEM image was recorded at $V_t = 1.8$ V, $I_t = 2$ nA. Topographic height range is 9 nm, and I_e ranges from about 0.2 to 2 pA; imaged area is 190×160 nm [137].

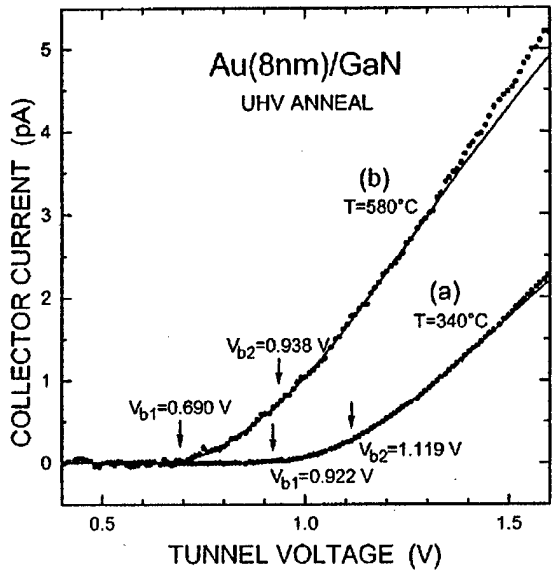


Fig. 35. BEEM spectra for two different Au (8 nm)/GaN samples. (a) Spectrum for a sample annealed in UHV at 340°C for 15 min. (b) Spectrum for a sample annealed at 580°C. Here transmission is increased, and barrier height is substantially decreased [19].

5.1.2.2. GaInP. In the past decade, spontaneous CuPt ordering of many III–V alloys has been widely observed in vapor phase growth on (001) substrates. In $Ga_{0.52}In_{0.48}P$ (written as $GaInP_2$ for simplicity), extensive theoretical [138,139] and experimental [140] work has been carried out to study the effect of the ordering-induced reduction of the crystal symmetry on the structural, optical and transport properties of the ordered material. Ordering induced changes in the band structure

of $GaInP_2$ are important for its application in advanced solar cells with very high conversion efficiency, for potential band-gap engineering and for fundamental studies of atomic ordering.

Perfectly ordered material has not been observed. The electronic states of a partially ordered structure can be interpolated from those of the totally disordered and the perfectly ordered semiconductor [139]. The degree of ordering is usually described by the ordering parameter η, where $E_g(\eta) = E_g(\eta = 0) - \eta^2 \Delta E_g(\eta = 1)$, with $\eta = 0$ and 1 to describe perfectly disordered and perfectly ordered material, respectively. $E_g(\eta = 0)$, the bandgap of disordered $GaInP_2$, is 2.01 eV at low temperature [141] and $\Delta E_g(\eta = 1)$, the maximum bandgap reduction for perfectly ordered $GaInP_2$, is 0.47 eV [142]. In the most ordered $GaInP_2$ structures, the highest ordering parameter $\eta \sim 0.6$ [143], and the degree of ordering is nonuniform on the local scale.

To characterize the structural and electronic properties of $GaInP_2$, we have applied the BEEM technique [15,16]. The undoped GaAs/$GaInP_2$ structures on n^+- and p^+- GaAs substrates were grown by MOCVD at $T_g = 650°C$. The structures consist of a 500 Å undoped GaAs buffer layer, a 1 μm $GaInP_2$ and a 50 Å GaAs cap layer. The details of the diode fabrication procedure have been published elsewhere [29]. A detailed analysis is presented here for the $GaInP_2$ samples grown on a (511) GaAs substrate and on a (001) GaAs substrate misoriented 6° toward $[111]_B$. From the low-temperature photoluminescence, the bandgap energy (after correction on the exciton binding energy) is ~2.00 eV for $GaInP_2$ grown on a (511) GaAs substrate and ~1.88 eV for $GaInP_2$ grown on a 6°$[111]_B$-(001) GaAs substrate. Therefore, we conclude that the $GaInP_2$ layer grown on a (511) GaAs substrate is highly disordered (hereafter, disordered $GaInP_2$) and the $GaInP_2$ grown on a (001) GaAs substrate misoriented 6° toward $[111]_B$ is highly ordered, $\eta \sim 0.5$ (hereafter, ordered $GaInP_2$) [139].

Fig. 36 shows room-temperature 1 μm × 1 μm STM images of the disordered (a) and ordered (b) $GaInP_2$ layer grown on n^+-GaAs substrates. One can see that the surface of ordered material forms [110]-oriented steps. For ordered $GaInP_2$ grown on misoriented substrates, the [110] steps are usually observed to form only single CuPt variant [144,145]. For disordered $GaInP_2$, the surface is found to be much flatter than that of ordered $GaInP_2$. The complementary BEEM images are also shown in Fig. 36. The observed contrast of the BEEM image for the ordered $GaInP_2$ sample is in direct correlation (antiphase) with the surface morphology, indicating high sensitivity and high spatial resolution of the BEEM technique.

To study the heterostructure transmission coefficient, we analyze the SD-BEEM spectra rather than the original BEEM spectra [16]. Fig. 37 shows the SD-BEEM spectra obtained from the experimental BEEM data by numerical differentiation with a 10 meV window. We associate two clearly pronounced features in the 77 K SD-BEEM spectra of disordered $GaInP_2$ with the Γ and L conduction minima contribution in $GaInP_2$. Theoretical fits to the SD-BEEM spectrum for the disordered $GaInP_2$ sample, using the MSIS model [5,85], are shown in Fig. 37 by the dashed line. The MSIS model fit describes the experimental BEEM spectrum reasonably well, giving ~90% probability of the electron scattering at the m–s interface, similar to our previous results for GaAs/AlGaAs structures [5]. The absence of a contribution from the X conduction minimum is due to strong X-electron attenuation in the GaAs cap layer [5].

The SD-BEEM spectrum of the ordered $GaInP_2$ sample presented in Fig. 37 show a very important difference from that of the disordered $GaInP_2$. Namely, we observe two high-energy peaks instead of one peak in disordered $GaInP_2$. We assign both high-energy peaks to be associated with the L valley contribution. In CuPt-type ordered $GaInP_2$, one of the four L valleys

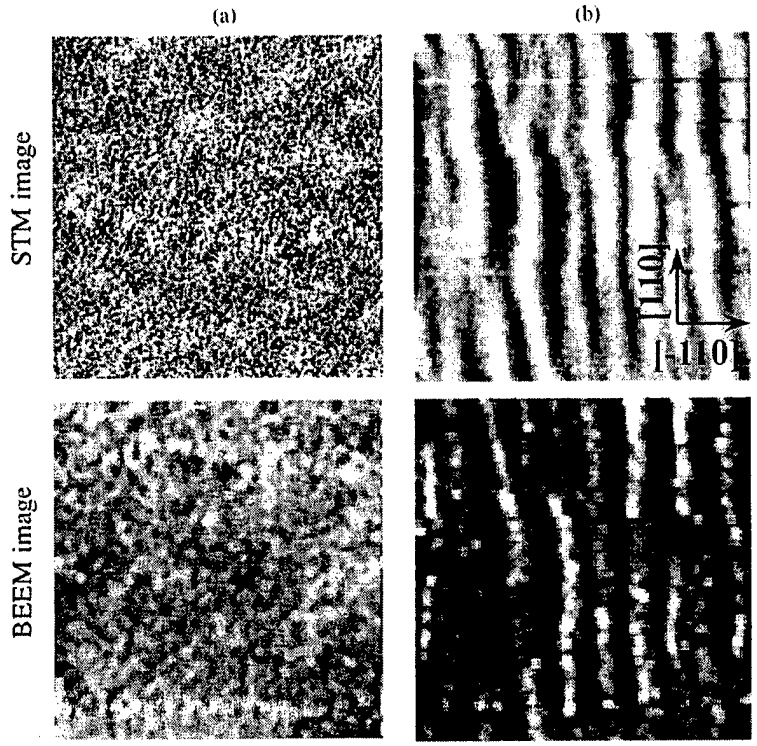


Fig. 36. Room-temperature 1 μm × 1 μm STM image (top) and BEEM image (bottom) of GaAs/GaInP$_2$ layer grown by MOCVD on (a) (5 1 1) n-GaAs substrate, (b) on n-GaAs substrate misoriented by 6° to $[111]_B$. A tip bias is − 1.7 V and a tunnel current is 4 nA [16].

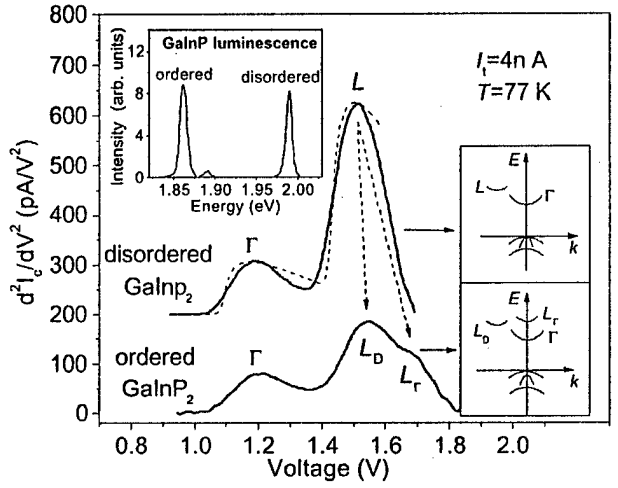


Fig. 37. SD-BEEM spectra for two GaAs/GaInP$_2$ (ordered and disordered) samples with 50 Å GaAs cap layers, taken at $T = 77$ K. For clarity, the spectra are shifted along the vertical axis. The MSIS model calculations are also presented (dashed line). The top inset shows the photoluminescence spectra for both ordered and disordered GaInP$_2$ samples, whereas two bottom insets show the schematic of the GaInP$_2$ band structure [16].

folds onto the $\bar{\Gamma}$ point (hereafter L_Γ), and the other three are folded onto the $\bar{D}$ point (hereafter L_D). A strong repulsion between the Γ valley and the folded L_Γ valley results in the bandgap reduction and in the increase of the Γ–L_Γ separation, while the energetic position of the L_D remains almost the same. Therefore, we conclude that the two observed high-energy peaks for the ordered sample are the contribution of the L valleys that are split due to ordering.

The "folded" zone-edge bands were observed experimentally in electro-reflectance [146] and Raman spectroscopy [147] measurements. In the recently reported electro-absorption experiments on ordered $GaInP_2$ ($\eta \approx 0.45$) [148], an additional feature was observed at $\sim$0.48 eV above the fundamental bandgap transition, and this feature was attributed to the back folded L conduction band. As the ordering decreases, this peak shifts to lower energies, with an asymptotic value of $\sim$0.33 eV above the fundamental bandgap transition for a totally disordered sample. In Refs. [146–148], due to the selection rules, only the contribution from the L valley folded onto the $\bar{\Gamma}$ point was observed. In contrast, we observe the contribution from all L valleys and, as a consequence, can measure directly the Γ–L separation in disordered $GaInP_2$ as well as the ordering-induced L valley splitting in ordered $GaInP_2$. According to our results, $\Delta(\Gamma\text{–}L) \cong 0.35$ eV for a disordered sample, $\Delta(\Gamma\text{–}L_\Gamma) \cong 0.47$ eV and $\Delta(L_\Gamma\text{–}L_D) \cong 0.13$ eV for an ordered sample. These results are in a good agreement with the theoretical predictions. Indeed, as pointed out by Zunger [149], it is possible to obtain the dependencies $\Delta(\Gamma\text{–}L_\Gamma) = \Delta(\Gamma\text{–}L)_{\eta=0} + 0.50\eta^2$ and $\Delta(L_\Gamma\text{–}L_D) = 0.42\eta^2$, using Table 1 and Fig. 3(b) of Ref. [150]. Then, taking $\eta = 0.5$ and $\Delta(\Gamma\text{–}L)_{\eta=0} = 0.35$ from our experiment, $\Delta(\Gamma\text{–}L\Gamma) \cong 0.475$ eV and $\Delta(L_\Gamma\text{–}L_D) \cong 0.11$ eV.

5.1.2.3. GaAsN. Recently, a new class of III–V alloys, where small amounts of nitrogen replace the group V-element, has attracted a great deal of attention because of the observation of a giant bandgap reduction. In the case of dilute $GaAs_{1-x}N_x$, for example, the bandgap is reduced by more than 0.4 eV at $x \sim 0.04$, indicative of a colossal bandgap bowing parameter [151–155]. Such a large tuning range suggests also a great potential for such alloys for a variety of optoelectronic applications [156,157].

To shed new light on the $GaAs_{1-x}N_x$ band structure we have applied the BEEM technique to study the electron transport in the conduction band of $GaAs_{1-x}N_x$ alloys [20]. A 1000 Å undoped $GaAs_{1-x}N_x$ layer and 1000 Å n^+-GaAs buffer layer were grown on n^+ (001)-oriented GaAs substrates by gas source molecular beam epitaxy at 420°C (the details of the growth conditions were published elsewhere [158]). A detailed analysis is presented here for a set of $Au/GaAs_{1-x}N_x/n^+$-GaAs with the nitrogen compositions of $x = 0, 0.003, 0.005, 0.007, 0.012, 0.017$ and 0.021. The composition of $GaAs_{1-x}N_x$ layers was determined from dynamic simulations of the X-ray diffraction spectra. High-resolution X-ray diffraction data also demonstrated that these layers are indeed single-phase alloys, in accordance with previous photoluminescence studies that have indicated the alloy formation already at $x > 0.001$–0.0025 [159,160]. To make the Schottky contacts, Au layers (65 Å thick) were deposited by thermal evaporation through a shadow mask at a background pressure of 2×10^{-7} Torr. The details of the diode fabrication procedure have been published elsewhere [29].

Fig. 38 shows the room-temperature BEEM spectra of $GaAs_{1-x}N_x$ for seven different nitrogen compositions. For all nitrogen compositions except $x = 0$, one can distinguish two thresholds in BEEM spectra. As the nitrogen concentration increases, the low-energy threshold shifts towards lower voltages whereas the high-energy threshold shifts towards higher voltages (the thresholds'

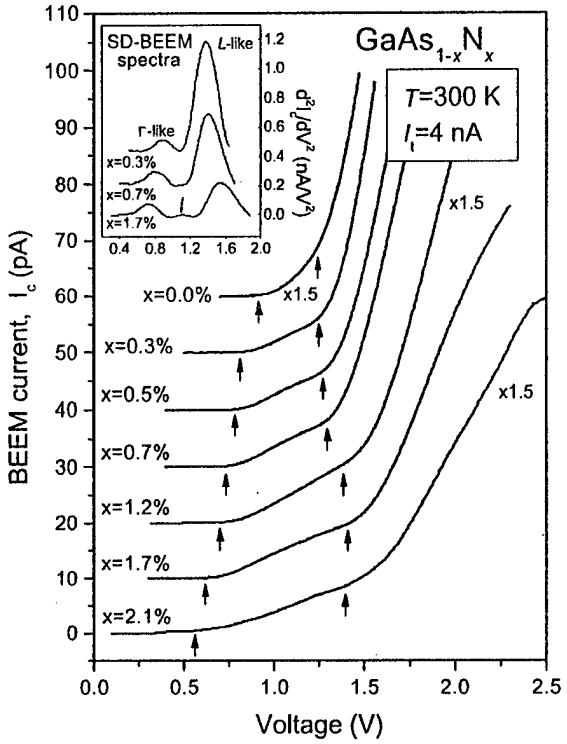


Fig. 38. Room-temperature BEEM spectra for seven different nitrogen compositions (from 0 to 0.021). For clarity, the BEEM spectra are shifted along the vertical axis. Arrows are eye-guides for the peaks position development. Insert shows the corresponding representative SD-BEEM spectra for three different nitrogen compositions (0.003, 0.007 and 0.017) [20].

development is shown by arrows in Fig. 38). This behavior is very different from our previously reported BEEM studies of AlGaAs [5] and $GaInP_2$ [16], where only the first threshold (counting from low voltages) can be unambiguously seen from the original BEEM spectra (all additional thresholds are visible to the naked eye only in the SD-BEEM spectra).

The representative room-temperature SD-BEEM spectra extracted from the experimental BEEM spectra by numerical differentiation with a 10 meV window are shown in the inset of Fig. 38. Two main features (peaks) observed in the SD-BEEM spectra we associate with the Γ-like and L-like conduction minima in $GaAs_{1-x}N_x$. Indeed, it is apparent from Fig. 38 that these peaks originate from gradually moving apart the Γ and L peaks of GaAs identified in our previous study [5]. The BEEM weighting of the different bands is proportional to their DOS's [85], and since nitrogen substitution results in the splitting of the fourfold L valley into the $a_1(L_{1c})$ singlet and $t_2(L_{1c})$ triplet states [161,162], the L-like band in the BEEM experiments is mostly weighted on the $t_2(L_{1c})$ triplet state. The SD-BEEM spectra of two $GaAs_{1-x}N_x$ samples, $x = 1.2$ and 1.7 (see inset of Fig. 38), reveal an additional weaker peak (indicated by the arrow), located $\sim$0.40 and $\sim$0.43 eV above the Γ-like state, respectively. (At lower nitrogen concentrations, this peak is probably masked due to the insufficient Γ-like and L-like peaks separation, whereas at higher nitrogen concentrations, due to the alloy-scattering-induced signal decrease, our sensitivity is too low to extract reliably the SD-BEEM spectra). This peak might represent the contribution from the

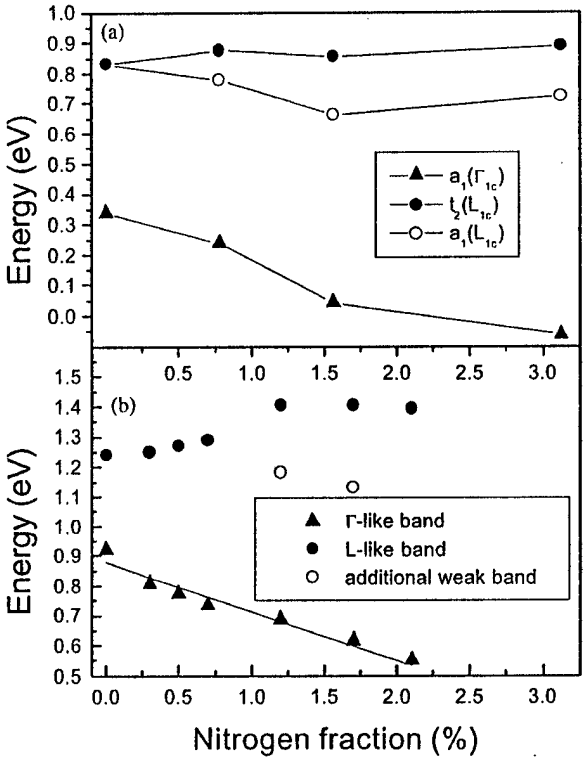


Fig. 39. (a) First-principles LDA calculations of the Γ_{1c} and L_{1c} compositional dependencies of $GaAs_{1-x}N_x$. (b) The compositional dependencies of the thresholds observed in the SD-BEEM spectra of $GaAs_{1-x}N_x$. The solid line is the best linear fit to the Γ-like threshold giving a slope of -16.4 eV [20,164].

$a_1(L_{1c})$ singlet state. The weaker amplitude of this peak matches the expected small DOS due to the increasing Γ-character of the $a_1(L_{1c})$ state in the alloy limit [163]. The compositional dependencies of the thresholds observed in the SD-BEEM spectra plotted in Fig. 39(b) are in a good agreement with recent first-principles theoretical calculations [164] presented in Fig. 39(a). The alternative L-like band identification as a localized resonant nitrogen level (see Ref. [165]) is very unlikely because in this case one would expect a resonant nitrogen-level contribution to increase with the nitrogen concentration (larger DOS), whereas the experimentally observed high-energy peak amplitude decreases with the nitrogen concentration. The absence of the X conduction minimum contribution in $GaAs_{1-x}N_x$ is due to the image potential-induced strong electron scattering in the spacer between the metallurgical m–s interface and the maximum of the barrier height [2], similar to our previously reported results [5].

As the nitrogen concentration increases, the $Au/GaAs_{1-x}N_x$ Schottky barrier (Γ-like threshold) decreases considerably, from ~ 0.92 eV at $x = 0$ down to ~ 0.55 eV at $x = 0.021$, as shown in Fig. 39(b). The solid line in Fig. 39(b) is the best linear fit to the experimental data (with a slope of -16.4 eV). Using $E_g(GaAs) = 1.42$ eV and $E_g(GaN) = 3.5$ eV at room temperature, the same slope of $E_g(GaAs_{1-x}N_x)$ would correspond to the bowing parameter of -18.9 eV. This value of the E_g bowing parameter is in a good agreement with the experimental estimates [152,166,167]. Thus, we conclude that the nitrogen-induced Schottky barrier reduction follows approximately the

bandgap reduction in $GaAs_{1-x}N_x$. This result, that is very important for device applications, indicates that the effect of the nitrogen incorporation on the valence band is small, in agreement with other studies [156,161].

5.2. *BEEM spectroscopy to study band offsets in buried semiconductor heterostructures*

To engineer useful heterojunction-based devices and predict their performance, the values of the band offsets must be accurately known. While originally invented to study Schottky barriers, BEEM spectroscopy can be effectively employed for the band offset measurements by comparative study of a reference metal/semiconductor system (e.g. Au/GaAs) and of an essentially similar system but with a buried semiconductor barrier beneath the m–s interface (e.g. Au/GaAs/$Al_xGa_{1-x}As$). In such a way, the band offset is obtained by subtracting the BEEM threshold of the reference sample from the BEEM threshold of the sample with the buried heterojunction barrier. Such structures have the advantage that the m–s (e.g. Au/GaAs) interface is invariant and allows a systematic comparison to be made as one varies the nature of heterostructure and alloy composition. In addition, BEEM offers certain advantages as compared to the standard techniques for the heterojunction band offset measurements, such as interband optical and macroscopic electrical techniques. Namely, BEEM allows independent measurements of the conduction band offset (ΔE_c) and the valence band (ΔE_v) offsets of "unbiased" heterojunctions as well as allows a high locality of the band-offset measurements. Spatially resolved BEEM spectroscopy is especially useful to study the band offsets of localized buried heterostructures, which are too small to contact by traditional methods (for example, see Section 5.4.3.2 for the BEEM study of buried GaSb quantum dots embedded into the GaAs matrix).

O'Shea et al. [29] have applied BEEM spectroscopy to measure band offsets in buried GaAs/$Al_xGa_{1-x}As$ heterojunction for $0 \leq x \leq 1$. For the single barrier case (see Fig. 40(a)), the $Al_xGa_{1-x}As$ barrier height was altered by systematically varying the Al composition. The barrier thickness was 100 Å and the GaAs capping layer was also 100 Å. Be δ-doping in the GaAs capping layer was used to flatten the conduction band. The conduction band offset first rises linearly to a maximum of ~ 0.35 eV at $x \sim 0.4$ and then falls gradually to ~ 0.3 eV at $x = 1.0$. This behavior is a consequence of the fact that $Al_xGa_{1-x}As$ is a direct semiconductor at $x \leq 0.42$ and an indirect one at $x \geq 0.42$. In the direct regime ($x \leq 0.42$), it was shown that the band offset of $\Delta E_c/\Delta E_g = 0.68$ measured by BEEM spectroscopy (see Fig. 40(c)) is in a good agreement with previous measurement by other techniques, thus demonstrating the accuracy of the BEEM spectroscopy technique. It was found that the conduction band-offsets are the same at $T = 300$ and 77 K. This result is in reasonable agreement with another BEEM study that was performed on a Au/GaAs/$Al_{0.2}Ga_{0.8}As$/GaAs single-barrier structure at $T = 7$ and 77 K [32], $Q_c = 0.58$ (0.60) at $T = 77$ K (7K). Note that, in addition to the band-offset measurements, the Au/GaAs/$Al_xGa_{1-x}As$ system was extensively studied for different aspects of multivalley electron transport in $Al_xGa_{1-x}As$ (Refs. [29,49,168]) as well as for quantitative analysis of the electron scattering at the m–s interface (see Section 2.4).

O'Shea et al. [169] have also employed BEEM to study the conduction band and valence band offsets in $GaInP_2$/GaAs single barrier heterostructures that are very promising to use in lasers, light-emitting diodes, heterojunction bipolar transistors, and solar cells. (To optimize these devices, control and understanding of the interface properties are essential. Yet, such fundamental

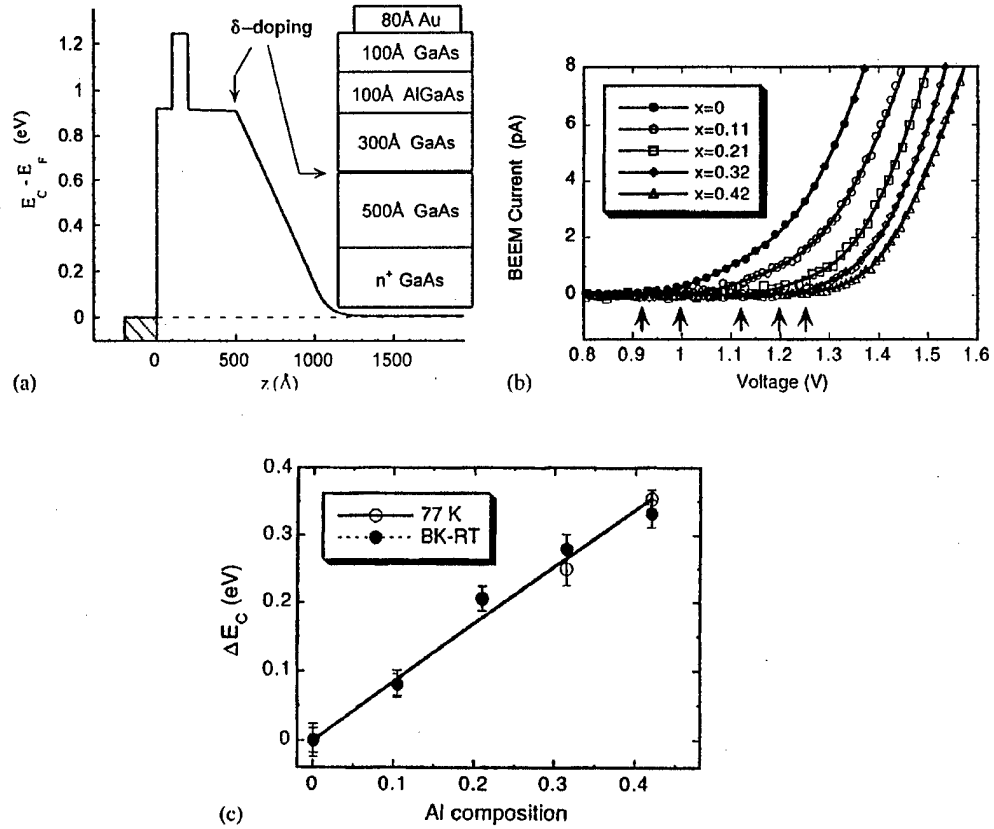


Fig. 40. (a) Calculated conduction band profile for $Al_xGa_{1-x}As$ ($x = 0.42$) single barrier structure assuming a Au/GaAs Schottky barrier of 0.92 eV. $T = 300$ K. The δ-doped sheet (Be, 1.1×10^{12} cm^{-2}) was used to flatten the bands. The MBE layer structure is shown in the inset. (b) Comparison of room-temperature (RT) BEEM spectra (point) for $Al_xGa_{1-x}As$ single barriers. Also shown are BK fits to the data (lines). Note the shift in the initial BEEM threshold (arrows) with Al composition gives the conduction-band offset. (c) RT and 77 K $Al_xGa_{1-x}As$ conduction-band offsets (points) measured by BEES. Linear fits (lines) at both temperatures give $\Delta E_c = (0.84\text{ eV})x$, or a fractional band offset of $Q_c = \Delta E_c/\Delta E_g = 0.68$. The linear curve fits and some data points are overlapping [29].

parameters as heterojunction band offsets are not well understood. Experimentally, values of conduction band offsets ranging from 30 to 390 meV have been reported [170]. Since the bandgap of $GaInP_2$ changes with the ordering parameter, it is likely that ordering is responsible for some of the variation in previous ΔE_c measurements.) The calculated band profiles and layer structures for the n- and p-type $GaInP_2$/GaAs single barrier samples grown by MOCVD at atmospheric pressure are given in the upper panel of Fig. 41. The δ-doping sheet concentration was calculated to cancel the band bending at room temperature near the Schottky barrier enabling a measurement of flat-band heterointerfaces. The GaAs buffer layers were grown at 650°C, and the susceptor temperature was ramped down to 610°C after the δ-doping for the $GaInP_2$ layer growth. BEEM spectra for the $GaInP_2$ single barriers on n- and p-type GaAs substrates are shown in the lower panel of Fig. 41. The BK model was used as a consistent procedure to determine the thresholds. Although the exact value of each threshold will depend on the model applied, the difference

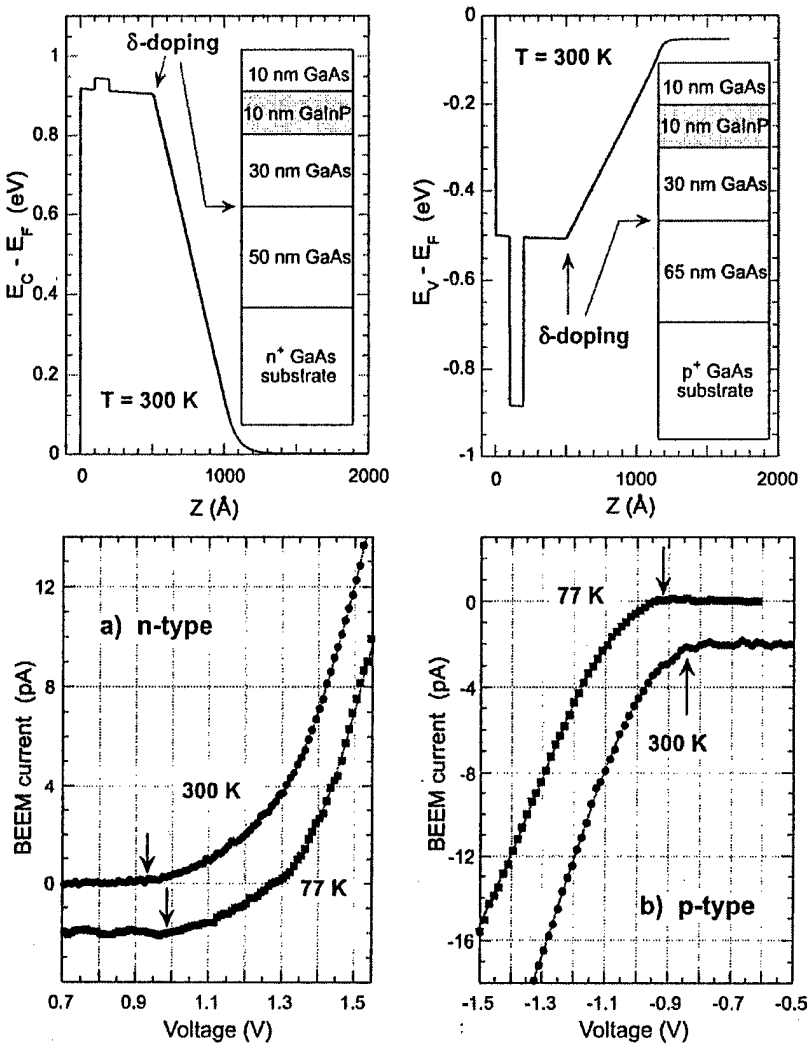


Fig. 41. Upper panel shows the calculated band profiles for n- and p-type $GaInP_2$ single barrier structures (the MOCVD layer structures are given in the inserts). Lower panel shows representative BEEM spectra (points) for (a) n- and (b) p-type samples containing $GaInP_2$ layers grown at 610°C. Lines are fits to the BK model to determine the thresholds, indicated by arrows. Curves offset by 2 pA for clarity [169].

between the BEEM thresholds of the single barrier sample and the reference sample should be model-independent. Thus, the band offsets are obtained by subtracting the BEEM threshold of the GaAs reference samples from the BEEM threshold of the $GaInP_2$/GaAs samples. $GaInP_2$/GaAs band offsets of 30 and 350 meV in the conduction and valence bands, respectively, were found at room temperature. Similarly, 77 K band offsets are 30 and 390 meV in the conduction and valence bands, respectively. Also, using the above BEEM results, it was possible to estimate the ordering parameter in the samples under study. Assuming that the sum of the n- and p-type BEEM thresholds should correspond to the $GaInP_2$ band gap, the BEEM measurements at

both 77 and 300 K give a bandgap reduction of ~0.1 eV compared to fully disordered GaInP, and this corresponds to an order parameter of $\eta \sim 0.45$.

The BEEM technique has been also applied to the Au/n-InAs/GaAs system to study the conduction band offset formation between InAs and GaAs as the thickness of the InAs was varied between 0 and 40 monolayers [132,133]. The band offset between InAs and GaAs was found to decrease with the InAs thickness. One monolayer of InAs lowers the barrier from ~0.9 (original Au/GaAs barrier) to ~0.8 eV. As the thickness of the InAs layer increases to three monolayers, the barrier was found to decrease further to ~0.74 eV. The barrier height remains approximately constant for up to 27 monolayers, and then it reduces to ~0.63 eV for the case of 33 monolayers (100 Å) and beyond. The detailed variation of the band offset was shown to correlate with the relaxation of the InAs layer, i.e. conduction band offset between strained InAs on GaAs depends on the strain, and it decreases as the strain relaxates through islanding.

Recently, BEEM spectroscopy has been successfully applied to determine the conduction band offset between AlInGaAs strained layer (active layer in 850 nm vertical cavity lasers) and a AlGaAs barrier located beneath the surface [30]. Also, BEEM spectroscopy has been used to study both the conduction and valence band offsets between InAs and AlAsSb at $T = 77$ and 300 K [31]. It was found that with the addition of As to AlSb, the conduction band offset between InAs and AlAsSb decreases despite the increase in the bandgap. The resulting decrease in the valence band position causes the InAs/AlAsSb band lineup to change from a staggered (type II) to a straddling (type I).

5.3. BEEM spectroscopy to study resonant semiconductor structures

5.3.1. Double barrier structures

The first BEEM study of resonant transport through a double barrier resonant tunneling structures (DBRTS) was conducted by Sajoto et al. [33] on GaAs/$Al_xGa_{1-x}As$ heterostructures buried beneath the Au/GaAs interface. Fig. 42 (left) shows the layer structure and calculated band profile of the DBRTS structure whose design included the δ-doped Be sheet to flatten the band profile at room temperature. A comparison of 77 K BEEM spectra for (a) the GaAs reference sample, (b) the DBRTS and (c) the 100 Å $Al_xGa_{1-x}As$ ($x = 42$) single barrier is given in the right panel of Fig. 42. The threshold for the DBRTS is clearly lower than that for the single barrier sample, which has the same barrier height, and the threshold position is in a good agreement with the calculated value of the quasi-bound state energy (~1.2 eV). At higher voltages, additional features are observed in the DBRTS spectra, and they were attributed to the opening of additional conduction channels associated with the higher-lying L and X valleys.

The above first study, where the temperature dependence of the BEEM spectra was used to investigate the hot electron transport, led to the further study of the electron-wave interference effect on the ballistic transmittance and device performance. Quantum-interference filters were designed and realized by using GaAs/$Al_{0.2}Ga_{0.8}As$ as a heterostructure of choice [171]. The designed devices are in essence the GaAs/$Al_{0.2}Ga_{0.8}As$ double barrier structures with one tunneling resonance (~50 meV above the Γ minimum of GaAs) and next above-barrier resonance (~200 meV above the Γ minimum of GaAs). The distance between $Al_{0.2}Ga_{0.8}As$ barriers of 9 ML thickness was 22 ML (11 ML) to produce a constructive (destructive) interference at the above-barrier quasi-bound state. It was shown that the SD-BEEM spectra accurately reproduce the transmittance functions of the designed half- and quarter-wavelength structures at 77 and 300 K

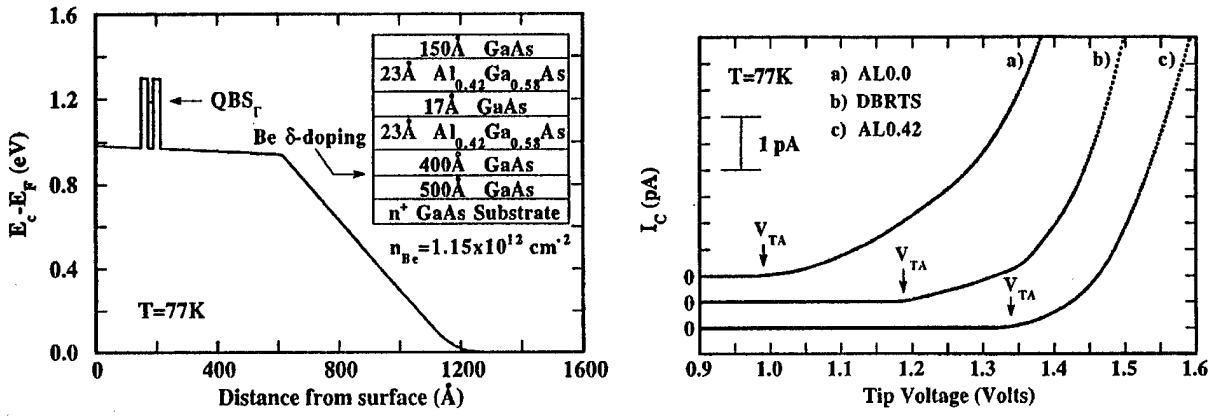


Fig. 42. Left panel shows the layer structure and calculated band profile of the DBRTS structure whose design included the δ-doped Be sheet to flatten the band profile at room temperature. Right panel shows the BEEM spectra at 77 K for (a) the reference GaAs sample, (b) DBRTS, and (c) GaAs/$Al_{0.42}Ga_{0.58}As$ single barrier sample [33].

(attaining nearly temperature-limited resolution) (see Fig. 43). The tunneling resonance was identified with a first peak (from low-energy side) in the SD-BEEM spectrum for both devices, whereas the above-barrier quasi-bound state position coincided with the higher-energy peak (dip) for constructive (destructive) interference. Recently, the 77 K SD-BEEM spectra of a half-wavelength filter were used to measure fluctuations in the thickness of buried GaAs and $Al_{0.2}Ga_{0.8}As$ layers [172]. By performing BEEM spectroscopy at several different spatial locations and by utilizing the relation between the thickness of the device and the energy of the quasi-bound states, single-monolayer fluctuations were detected in the thickness of both the GaAs quantum well and the surrounding $Al_{0.2}Ga_{0.8}As$ barriers.

5.3.2. *Superlattices*

A miniband in a short period superlattice (SL) covers a rather broad energy range compared to a double barrier resonant tunneling diode and, therefore, electron transport through its states is expected to be more pronounced and better resolved.

Cheng et al. [135] have applied BEEM to study electron transport in InAs/AlSb SLs. The barrier was too low for the SL to allow consistent probing by BEEM spectroscopy. However, the SL BEEM signal was elevated above the background noise after repeated stressing of the metal surface. A BEEM threshold of 0.8 eV was observed for the Au/24 Å-period SL system after the stress treatment.

Smoliner et al. [35,173] have used a buried 10-period 30 Å GaAs/25 Å $Al_{0.4}Ga_{0.6}As$ SL (with one below-barrier miniband) on Au/GaAs Schottky diodes to study the energetic current distribution in BEEM at $T = 100$ and 300 K (GaAs cap layer is 300 Å). The authors showed that the miniband results in a BEEM current threshold clearly below the height of the $Al_xGa_{1-x}As$ barriers, as shown in Fig. 44. In contrast to the Au/GaAs reference sample, where the BEEM current follows the 2 or 5/2 power law, the SL spectrum is almost linear up to a sample bias of 1.3 eV (above this bias, electrons start to overcome the AlGaAs barriers). The theoretical calculations, based on the SK model (with the incorporation of an additional SL transmission coefficient

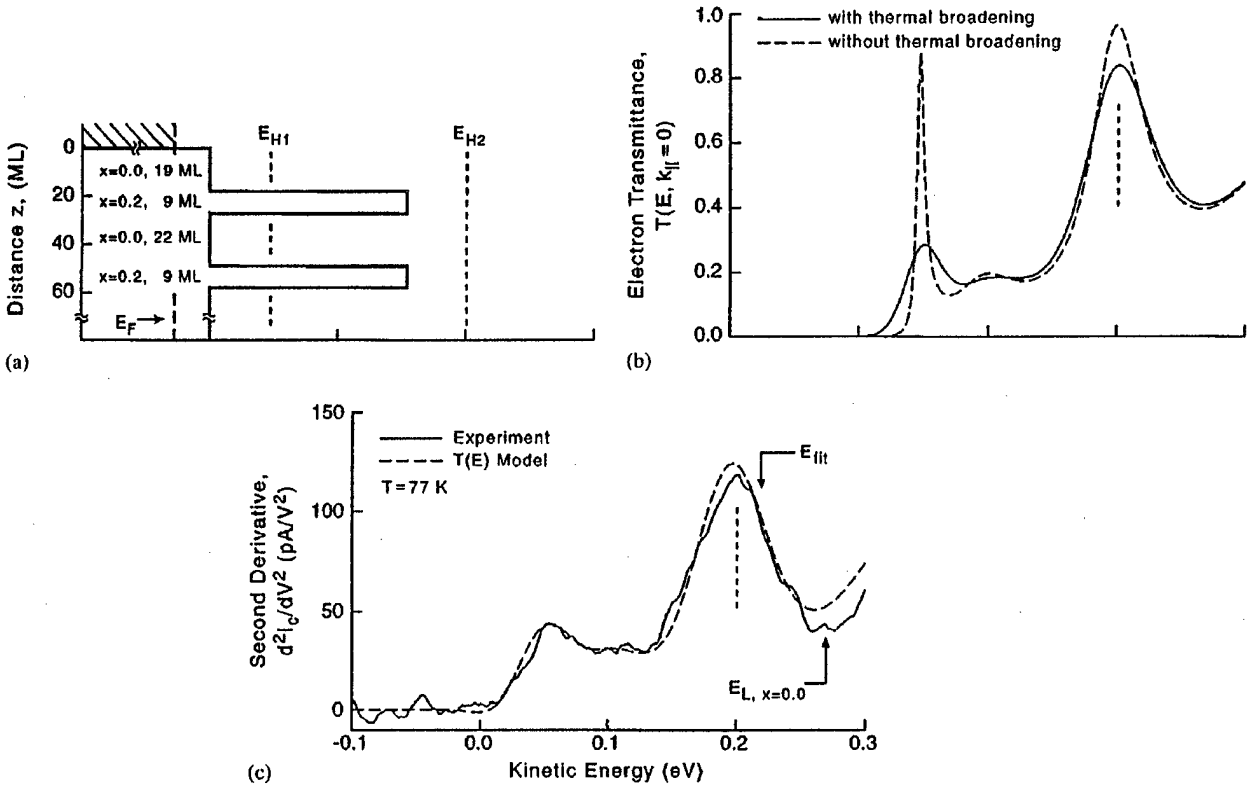


Fig. 43. (a) The band diagram and (b) the corresponding normal-incidence electron transmittance function $t(E, k_{||} = 0)$ with and without thermal broadening for the designed half-electron-wavelength resonant device which has one tunneling resonance E_{H1}, and one above-all-barriers resonance E_{H2}. (c) Second derivatives of the experimental spectrum and of the BEEM $t(E)$ model spectrum for data acquired at $T = 77$ K. The spectrum accurately reproduces the shape of the thermally broadened $t(E, k_{||} = 0)$, which represents the resolution limit of BEEM spectrum [171].

into it) are in a good agreement with the experiment and confirm the linear BEEM spectral shape in the SL energetic range. It was shown that it is an electron refraction at the Au/GaAs interface (taking into account the $k_{||}$ conservation and difference in the electron effective masses) that leads to the almost inversion of the hot electron distribution (in Au, the electron distribution is peaked toward the Fermi energy of the STM tip, but in GaAs the electron distribution is peaked toward the Schottky barrier). Thus, as the tip-to-base voltage increases, an increasing number of electrons are available at the SL energy, and the BEEM current increases continuously, leading to the almost linear dependence on the tip-to-base bias. However, if the resonant structures are located directly below the metal base, the authors observed a step-like behavior of the BEEM spectrum that is explained by direct coupling (tunneling) of ballistic electrons in the Au base to the resonant level underneath [174]. It was also argued that assuming $k_{||}$ conservation, the direct tunneling into the resonant state, when there is no intermediate refraction at the Au/GaAs interface, would provide effective filtering for both $k_{\perp}$ and $k_{||}$. Thus one would expect better spatial resolution of BEEM. However, there are no experimental data confirming this prediction.

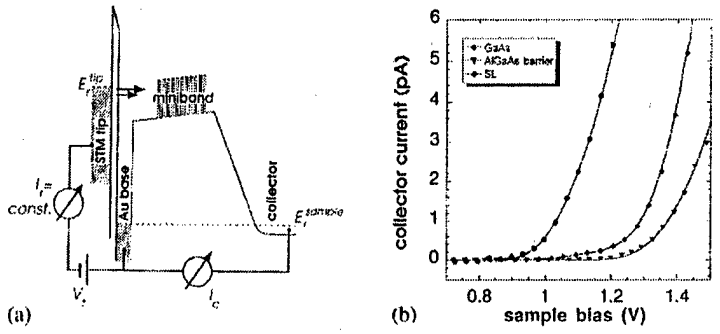


Fig. 44. (a) BEEM setup and conduction band profile with miniband regime of the superlattice sample. For a high enough bias V_t, carriers can be injected into the subsurface semiconductor structure and will be collected at the highly doped collector region. No external bias is applied across the semiconductor. (b) BEEM spectra obtained at room temperature on a sample with a conventional Au/GaAs(3 1 1)A Schottky barrier, a 100 Å thick AlGaAs barrier buried 300 Å below the interface, and on a superlattice sample. Solid curves show thermally broadened power-law fits for the GaAs and the AlGaAs sample, and a weighted fit to the data of the superlattice sample [173].

Heer et al. [34] showed that for biased 10-period 30 Å GaAs/25 Å $Al_{0.4}Ga_{0.6}As$ SL (buried 300 Å below the Au/GaAs interface), both the position of and the transmission of the miniband are a function of the applied base-to-collector bias. In the SD-BEEM spectrum numerically extracted from the original BEEM spectrum, the miniband manifests itself as clear peak. Biasing the SL results in a shift of the miniband peak position in the SD-BEEM spectra, in accord with the transmission calculations based on the SK model. In addition, the BEEM current magnitude is reduced as soon as a base-to-collector bias is applied, and it was explained in terms of the bias-dependent energy width of the miniband. This result is in accord with hot electron spectroscopy results, that in biased superlattice, a significant decrease of the miniband transmission with increasing electric field occurs due to the localization of the electron wave function [175]. As useful application of this experiment, the GaAs/AlGaAs SL can be employed as energy filter to study the energetic distribution of ballistic electrons in BEEM experiments by deconvolution of the transmitted electron current with the calculated miniband transmission as a function of collector bias.

5.4. BEEM imaging to study defects and nanostructures

5.4.1. Study of buried defects and dislocations

One example of the buried dislocation imaging in BEEM experiments was presented in Section 3.3. for the $In_xGa_{1-x}As$/GaAs interface buried 400–700 Å below the m–s interface. For silicide–silicon systems, in-situ BEEM technique was successfully applied by Fernandez et al. [72,73] ($NiSi_2$/Si at room temperature) and Sirringhaus et al. [78,80–82,176] ($CoSi_2$/Si at 77 K) to investigate the contrast mechanisms by which individual interfacial defects (misfit dislocations and point defects) can be resolved. It was found that in $NiSi_2$/n-Si(111) and $CoSi_2$/n-Si(111) systems, misfit dislocations and point defects buried ~25–30 Å beneath the surface locally enhance the scattering probability at the interface, as shown in Fig. 45. The increase of the BEEM current at the dislocation was explained by increased electron scattering at the dislocation core. The scattering broadens the $k_{||}$ momentum distribution, and thus facilitates the electron transmission into the

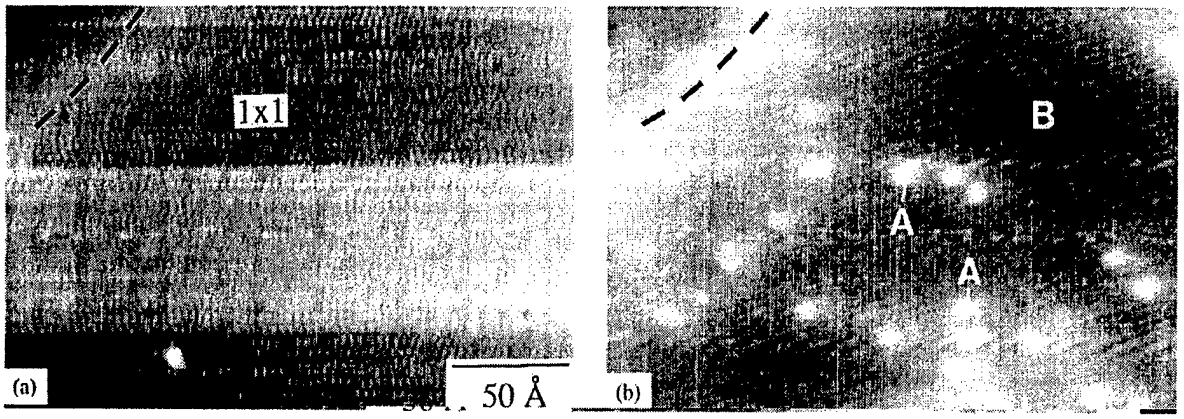


Fig. 45. STM topography (a) and BEEM image (b) of a 32 Å $CoSi_2$/n-Si(1 1 1) film. The corrugation in the STM topography image is not due to the unreconstructed (1 × 1) atomic surface structure but due to small mechanical vibrations. The BEEM current is enhanced at individual interfacial dislocations (dashed line) and subsurface point defects (A) due to hot carrier scattering. (V_t = 1.8 V, I_t = 10 nA) [81].

off-axis conduction bands of Si(111). The sharp BEEM dislocation profiles (in favorable cases a spatial resolution below 10 Å has been obtained [77,81]) and observation of the quantum size effect [79] imply that the electron transport in the metal is ballistic, and, therefore, most of the electrons reach the interface with $k_{||} \sim 0$. Increased BEEM current at dislocations indicates that the $k_{||}$ is mostly conserved in dislocation-free regions, although to some extent the probability for scattering exists everywhere at the epitaxial $CoSi_2$/Si interface [81]. Assuming $k_{||}$ conservation, the electron scattering at interfacial defects should increase the BEEM current in the case of $CoSi_2$/n-Si(111) and decrease the BEEM current in the case of $CoSi_2$/p-Si(111) (where holes are transmitted into the zone centered valence band minimum), in accordance with the experiment [77,80]. The same mechanism was invoked to explain the enhancement of the BEEM current observed in the vicinity of structural defects in the $NiSi_2$/Si(111) [72]. Note that no variations of the Schottky barrier have been observed at the silicide/Si(111) interface. In contrast, on $CoSi_2$/Si(100), certain interface dislocations and other defects lower the Schottky barrier by up to 0.1 eV on a nanometer scale [81]. In the future study, the observed Schottky barrier reduction will help to understand the effect of the presence of such defects on the average Schottky barrier height extracted from the transport characteristics of a macroscopic diode.

5.4.2. Study of lateral quantum structures

The laterally patterned quantum wires fabricated on modulation-doped GaAs/AlGaAs heterostructures were studied by BEEM at T = 4.2 and 300 K [37,46,177]. In contrast to bulk samples, the collector electrode consists of a 2D-electron layer at the AlGaAs/GaAs interface. After evaporation of ~70 Å Au film, wires were directly observed both in STM and BEEM images, as shown in Fig. 46. At T = 4.2 K, no decrease of STM resolution was found in comparison with room temperature measurements. The on-wire BEEM current was characterized by increased magnitude due to the lowering of the Schottky barrier (in contrast to the off-wire case, where the Schottky barrier is defined by $Al_xGa_{1-x}As$ ($x \sim 0.3$), the BEEM onset for the on-wire case is defined by the Au/GaAs Schottky barrier).

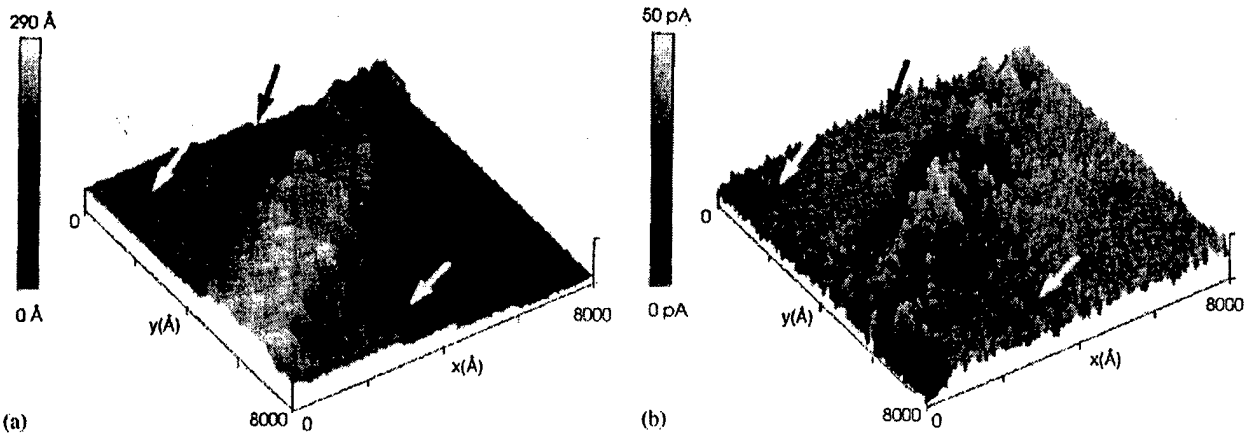


Fig. 46. (a) 3D topographic image of a single quantum wire ($V_t = 1.8$ V, $I_t = 5$ nA). (b) Simultaneously measured BEEM current image. The arrows indicate accumulated Au clusters in the etched region [37].

Westwood et al. [178] had applied BEEM to study GaAs/AlAs buried lateral period SLs (100 Å GaAs cap layer). While the STM image shows the usual grained surface due to the Au clustering during deposition, the BEEM image is dominated by striations aligned with the surface steps. A spectroscopic examination of the potential barrier across the sample surface reveals a bimodal distribution consistent with a modulation of $Al_xGa_{1-x}As$ mole fraction perpendicular with the surface steps.

5.4.3. *Study of single quantum dots*

5.4.3.1. InAs quantum dots. The first BEEM study of buried single quantum dots (QDs) was conducted on InAs self-assembled QDs buried spatially beneath a Au/GaAs interface [38]. For BEEM experiments, the InAs dots were grown on top of a 300 Å undoped GaAs buffer layer and covered with 50–85 Å GaAs cap layer. Fig. 47 shows a 7500 Å × 7500 Å STM image taken with a 1 nA tunnel current which shows the surface features above several dots covered by a 50 Å GaAs cap layer. The features are ~1000 Å in diameter and 30–50 Å high (the dip near the center of each dot was attributed to the strain-induced preferential buildup of the cap GaAs layer away from the center of the dot during the growth process). Gold grains, with diameter ~200 Å are also visible. Fig. 48 shows a high-resolution room-temperature STM image of a single InAs QD, capped with a 75 Å GaAs layer. Also shown is the corresponding BEEM image taken with $I_t = 2$ nA and $V_t = 1.5$ V, well above the Schottky barrier height. A strong enhancement of the BEEM current is observed on the QD buried beneath the surface. To estimate quantitatively the BEEM transport across the QDs, the BEEM spectra were taken on and off a single dot (right window in Fig. 48). The measured current spectra exhibit fine structure consistent with resonant tunneling through two 0D states of the dot (with energetic separation of ~0.1 eV). The BEEM images and spectra clearly show the power of BEEM to probe energetic states of individual, nanometer scale, semiconductor quantum structures buried beneath the surface.

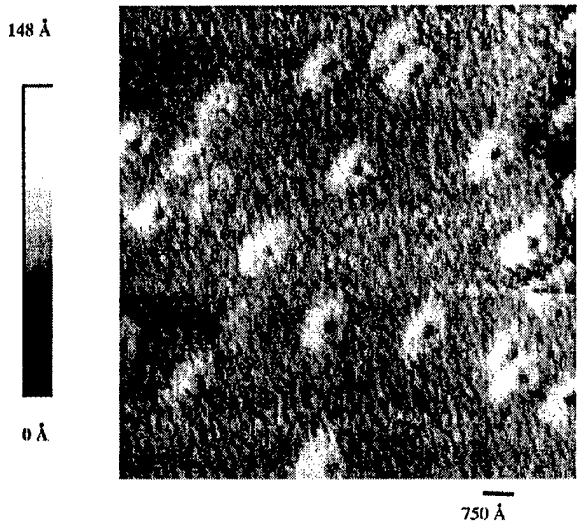


Fig. 47. STM image of surface features in a region where several InAs quantum dots are present and capped with a 50 Å GaAs layer and 85 Å Au layer. The dips near the center of each feature represent the positions of the dots beneath the surface [38].

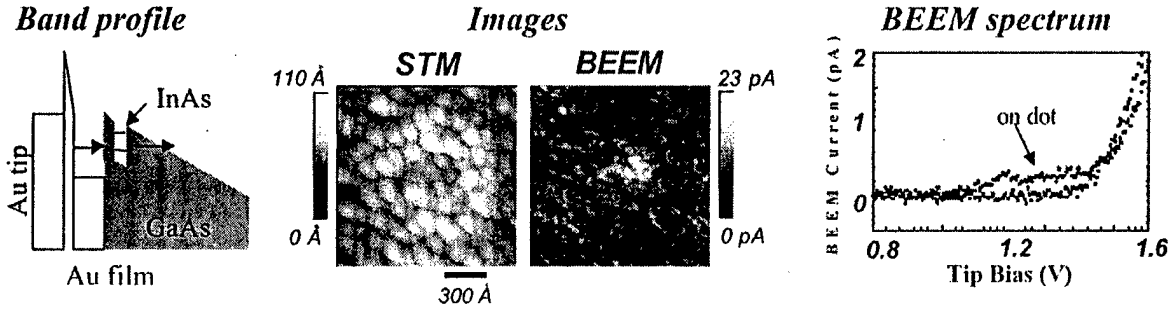


Fig. 48. (from left to right) Schematic of band profile, STM and BEEM images, and characteristic BEEM spectra for InAs/GaAs self-assembled single quantum dots [38].

5.4.3.2. GaSb quantum dots grown on GaAs. GaSb self assembled QDs grown by MBE on GaAs exhibit a staggered (type II) band lineup with a potential barrier in the conduction band [179–181]. Therefore, electron–hole recombination in these structures is spatially indirect, so that optical measurements do not provide adequate information about the conduction band offset between the GaSb dots and GaAs. Along with that, traditional transport methods cannot measure this local band offset because of the small ($\sim$500 Å) lateral dot size, and any offset measurement would average over areas with and without dots. In BEEM, however, carriers are locally injected into semiconductor structures to spectroscopically probe buried interfaces on a nanometer scale.

Fig. 49 shows STM and BEEM images of a single, GaAs capped, GaSb dot. In the STM image, a roughly circular feature $\sim$500 Å in diameter and $\sim$50 Å tall, marks the lateral position of the buried dot. The area in the BEEM image aligned with the dot profile in STM is darker than the surrounding region, implying that the BEEM current through the dot is reduced due to electron

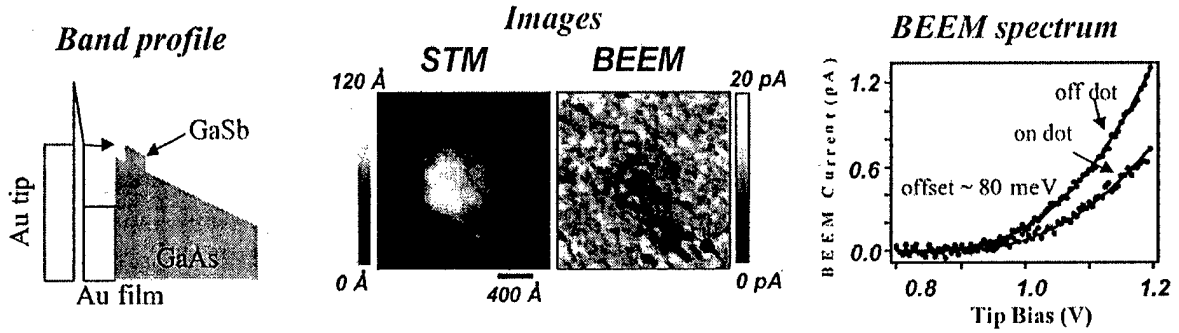


Fig. 49. (from left to right) Schematic of band profile, STM and BEEM images, and characteristic BEEM spectra for GaSb/GaAs self-assembled single quantum dots [39].

reflection off the dot's potential barrier. The height of this barrier, i.e. the local band offset, can be extracted from the changes in BEEM spectra between the on and off cases. The on-dot and off-dot BEEM spectra of several dots were fitted by using a modified Bell-Kaiser planar tunneling model [2,58], giving a local conduction band offset for GaSb dots on GaAs of 0.08 ± 0.02 eV. For details of this study, see Ref. [39].

5.4.3.3. Phosphide-related quantum dots. Phosphide materials are widely used to fabricate light-emitting diodes and injection lasers emitting in the visible spectrum. The BEEM study of the properties of pseudomorphic $Al_{0.3}In_{0.7}P$ QDs grown by low-pressure MOCVD on GaP was recently reported in Ref. [40]. AlInP self-assembled QDs were grown on top of a thick undoped GaP buffer layer (grown on the n^+-GaP substrate) and covered with 50 Å GaP cap layer. Finally, a 60 Å Au film was thermally evaporated to form a good base layer for BEEM measurements. From the STM images, scanned at room temperature with a tunneling current of 4 nA, a characteristic QD was estimated to be approximately 700 Å width and 50 Å of height. In simultaneously taken BEEM images, with the tip bias of 1.5 V, the self-assembled QD-induced contrast was observed, with a reduced BEEM current for the on-dot electron injection. Fig. 50(a) displays the spectroscopic data that are obtained when the STM tip is positioned on and off a single $Al_{0.3}In_{0.7}P$ QD. The solid line represents the theoretical fit to the experimental data using the BK model. When the tip is away from the dot, it essentially measures the barrier height of the Au/GaP interface, which is determined to be 1.27 ± 0.05 eV. However, when the tip is on the dot, the situation is quite different. The ballistically emitted electron will experience two different interfaces, which are Au/GaP and GaP/$Al_{0.3}In_{0.7}P$. Since the threshold for the BEEM current on the dot is approximately measured as 1.4 ± 0.05 eV, a local conduction-band offset between GaP and $Al_{0.3}In_{0.7}P$ can thus be determined to be 0.13 ± 0.1 eV.

Another example is the BEEM study of the self-assembled InP QDs on $Al_{0.5}In_{0.5}P$ (50 Å) grown on a GaAs matrix [182]. The dots are covered with $Al_{0.5}In_{0.5}P$ (50 Å) and finally capped with 50 Å of GaAs. Thus the InP quantum dots are sandwiched in a double barrier heterostructure of GaAs/$Al_{0.5}In_{0.5}P$/InP-QD/$Al_{0.5}In_{0.5}P$/GaAs. The BEEM spectra taken while the tip is away from the dot, as shown in Fig. 50(b), reveal the presence of a single threshold at 1.44 eV corresponding to the charge transport over the $Al_{0.5}In_{0.5}P$ barrier. However, when the tip is positioned on the QD, an additional threshold is observed at 1.12 eV, which is attributed to the

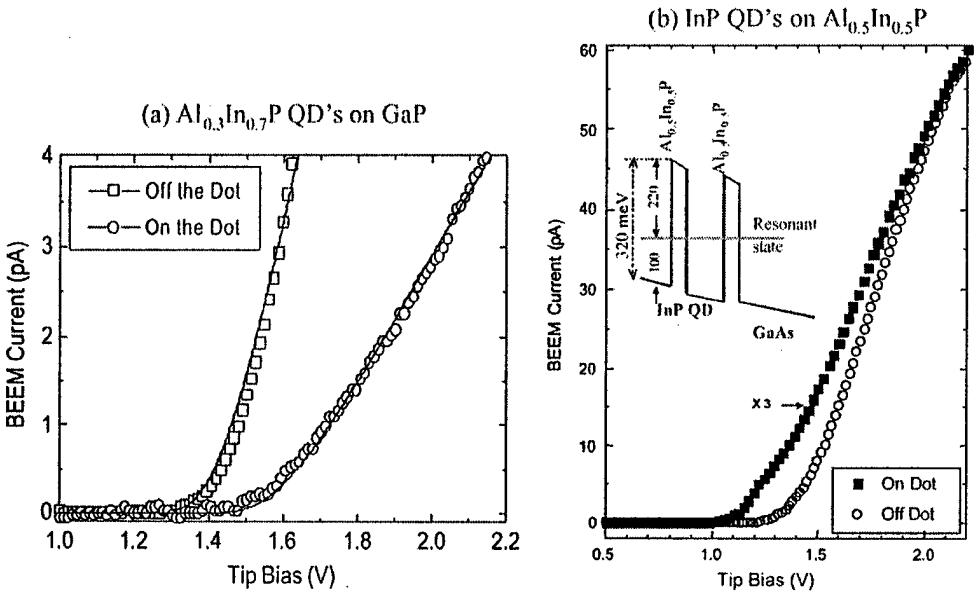


Fig. 50. BEEM current spectra taken on and off the QD for the case of (a) $Al_{0.3}In_{0.7}P$ QDs on GaP and (b) InP QDs on $Al_{0.5}In_{0.5}P$ grown on a GaAs matrix. A band diagram is shown in the inset [40,182].

resonant state in the double barrier heterostructure. The relative intensity of the BEEM current magnitudes above the thresholds further suggests that the transmission through the resonant state is significantly higher than over the barrier itself. By taking 0.92 eV as the Schottky barrier on GaAs, the band offset for GaAs/$Al_{0.5}In_{0.5}P$/GaAs is determined as 1.24 eV. Further, assuming the band offset for InP-QD/$Al_{0.5}In_{0.5}P$ barrier as 0.34, the relative position of the resonant state is determined as 0.10 ± 0.02 eV with respect to the GaAs conduction band minimum. A band diagram thus constructed is shown as an inset in Fig. 50(b).

5.4.4. BEEM study of magnetic structures

The spin-dependent electron scattering behavior of thin ferromagnetic films and the spin-filtering effects on transport across ferromagnetic/nonferromagnetic metal interfaces are fundamental to the understanding and application of the magnetotransport properties of magnetic multilayered systems. With the ever-shrinking size of magnetic devices and the promise of novel magnetic behavior there is a strong need to study magnetism at smaller and smaller scales, below the resolution limit of standard magnetic imaging techniques available today.

Following the initial study of a magnetic multilayer by BEEM spectroscopy [183], a variation of BEEM has been developed to image the magnetic structure in thin-film multilayers with nanometer resolution [97,184]. In the nominally uncoupled Co/Cu/Co (25 Å/45 Å/12 Å) trilayer system magnetic domains are found to occur on a $\sim$0.5 μm-length scale and less. The Co/Cu/Co trilayer, a ferromagnetic-normal-metal-ferromagnetic system that is the basic component of spin-valve and GMR devices, was grown on Si(111) substrate precoated with a thin Cu/Au bilayer to form a high-quality Au/Si Schottky barrier interface, and a Cu seed layer for the Co growth.

When the unpolarized current from the STM is incident upon the top ferromagnetic layer, the spin-dependent scattering results in strong attenuation of one of the spin components. When

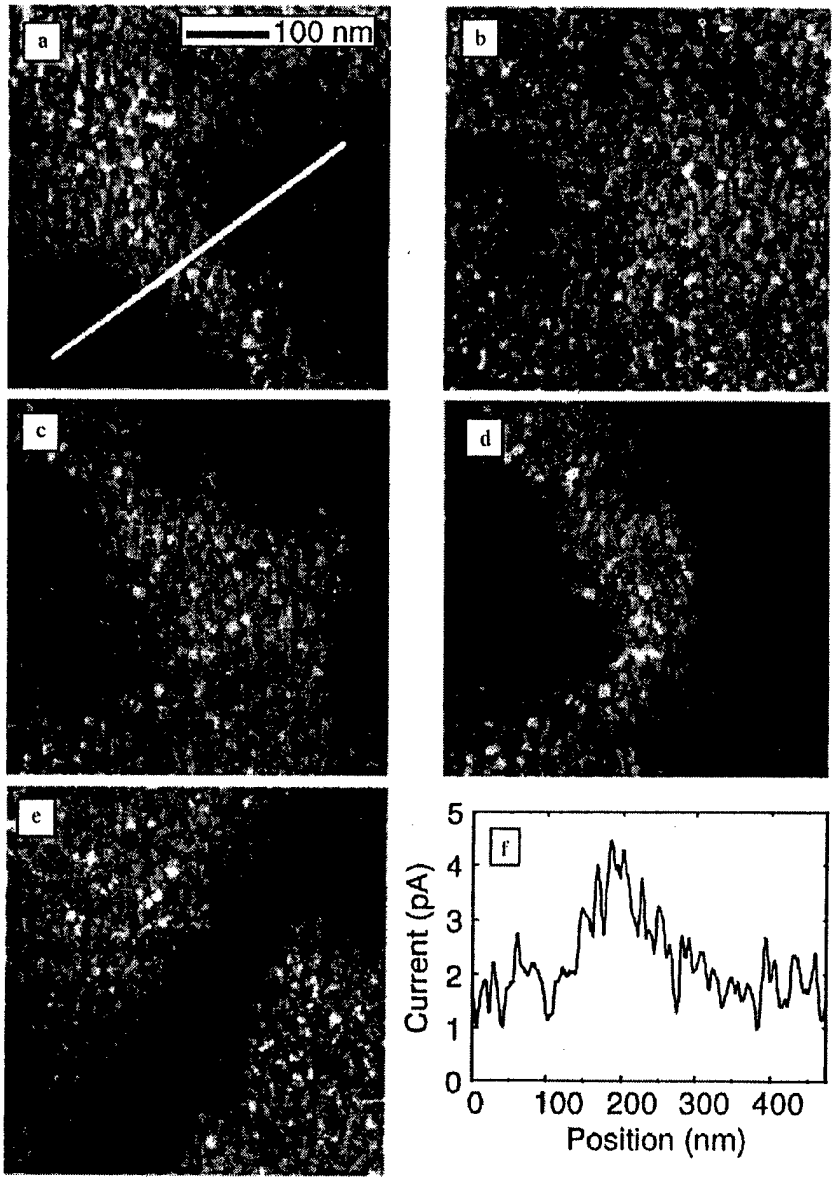


Fig. 51. (a)–(e) BEEM images (500×500 nm^2) taken at a fixed position in a varying magnetic field applied parallel to the film plane, (a) before the applied field, and in applied fields of (b) -40 Oe, (c) 10 Oe, (d) 20 Oe, and (e) 40 Oe. The linear gray-scale range for each is from 1.0 pA (black) to 5.0 pA (white). $V_t = 1.5$ V and $I_t = 2$ nA. (f) is a cross-sectional view of the position indicated in (a) [184].

the magnetization directions of the two layers are aligned, only one of the spin components will be heavily scattered in passing through them, yielding a comparatively high BEEM current. In the case of misalignment between ferromagnetic layers, both components will be strongly attenuated. This process is very analogous to "polarizer-analyzer" experiments in optics. It was found that by applying magnetic field $\sim$40 Oe, the BEEM contrast disappears indicating a saturated state of magnetic alignment, when the parallel alignment of two layers occurs [97,184], as shown in Fig. 51.

It was shown also that complete antialignment is never obtained over large, > 1 μm, areas for any values of magnetic field, indicating some residual coupling between the Co layers.

By hot electron (1–2 eV) BEEM spectroscopy taken as a function of Co layer thickness and relative magnetic alignment, the spin-dependent attenuation length of thin Co films and the energy dependence of the relative transmission factors for tunneling-injected electrons with different spin orientation were estimated [97]. The electron attenuation length in Co was estimated to be ~ 8 Å for electrons with spin orientation opposite to the Co layer magnetization, and ~ 21 Å for their parallel orientation, that are much smaller than the hot electron attenuation lengths found in free electron metals [86].

The filtering effect of the Co/Cu interface was demonstrated by studying the Co/Cu/Co trilayers by BEEM [97]. The strong reduction of the BEEM current when ~ 2–3 Å of Co is evaporated on top of Cu/Au/Si system was attributed to the band structure mismatch at the Co/Cu interface [Cu(111) has no propagating momentum states in a much larger cone around the film normal than Co(111)].

6. Metal–insulator–semiconductor systems

Study of electron transport in insulators (e.g. oxides) is very demanding for MIS (MOS) device performance and reliability. In the presence of an additional insulator layer between the metal and semiconductor, the distribution of electrons entering the semiconductor is modified by the insulator band structure as well as by electron–phonon scattering in the insulator layer. The electronic charge trapping and breakdown characteristics, which are determined by defects and impurities in the bulk insulator or at its interfaces, become increasingly sensitive to local fluctuations in material properties as device dimensions are reduced. The standard techniques (such as Fowler–Nordheim tunneling injection and internal photoemission [185–187]) to characterize the degradation mechanisms are still macroscopic (0.01–1 mm^2) and require high fields across the oxide to energize the injected electrons.

So far two insulators, SiO_2 and CaF_2, were extensively studied in BEEM experiments. SiO_2 is a material of natural choice for today's Si-based MOS devices. It is a chemically stable dielectric, it has excellent passivating and dielectric properties and its quality can be consistently controlled during production [188]. The performance and reliability of modern MOS devices is ultimately dependent on the electronic properties of SiO_2. Although both intrinsic and extrinsic electronic properties of SiO_2 have been the subject of extensive studies by using macroscopic techniques [185], the ongoing minimization of electronic devices requires the atomic-scale control and characterization of the material properties. CaF_2, in its turn, is an excellent choice for an epitaxial insulating layer in silicon-based MIS structures due to its large bandgap (12.1 eV) and its small lattice mismatch with Si (0.6%) allowing a growth of ultrathin flat CaF_2 layers as the insulating film in MIS devices. Si segregation to the surface, as observed in case of Au/Si interface, is strongly reduced by the CaF_2 intralayer [25]. Epitaxial CaF_2 films on Si(111), grown at elevated temperatures under UHV conditions, are considered to be promising candidates for replacement of SiO_2 in MOS technology. Because of the layer-to-layer growth of CaF_2, the more homogeneous insulator film can be grown with less density of interface states at the epitaxial CaF_2/Si(111) interface [189].

Cuberes et al. [25] first used the BEEM technique as a new, microscopic method to study electronic properties of very thin CaF_2 insulating films, and later BEEM was applied to study SiO_2 properties in MOS devices [26,64]. In this section we will review the main BEEM results for metal/SiO_2/Si and metal/CaF_2/Si systems.

6.1. Insulator conduction band effect

6.1.1. Metal/SiO_2/Si MOS system

Ludeke et al. [26,190] observed a BEEM current threshold near $V_t = 3.9$ V in 32 Å Pt/(27-62 Å) SiO_2/Si(100) structure, and it was interpreted in terms of current transport through the SiO_2 conduction band. The same threshold position was observed later by Kaczer et al. [191] in 40 Å Pt/27 Å SiO_2/Si(100) system. In the BEEM imaging mode, the first detection of high transmission regions in thin oxide was presented [190]. These were characterized by relatively small 1–2 nm patches that exhibited much lower thresholds ($\sim$1.1 V) than those observed for normal oxides ($\sim$4 V). It was explained in terms of defect-related or direct tunneling through the oxide bandgap.

In the case of thin SiO_2 (30 Å), the very well-defined thresholds (with essentially zero BEEM current below the threshold) were observed for Au/SiO_2/n-Si(100) (4 eV) and Ir/SiO_2/n-Si(100) (5.6 eV) [192], as shown in Fig. 52. In the case of ultrathin SiO_2 (10 Å), a different behavior was observed. For Ir/SiO_2/Si, a much lower threshold was observed at 1.5 eV. Since this threshold is different from the Schottky barrier at the Ir/Si interface (0.9 eV) and the signal is strong, it was explained not by direct tunneling through the barrier, but through a possible resonant state in the SiO_2 layer.

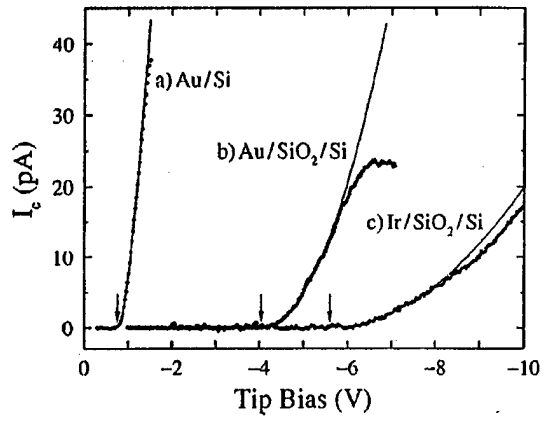
The oxide contribution to the BEEM transport was obtained by subtraction of the BEEM spectrum with zero SiO_2 thickness from the spectrum with finite SiO_2 thickness, after its correction for the electron impact ionization in Si (see Fig. 53). The authors studied the derived-from-experiment transmission probabilities across the SiO_2 layer as a function of tip-to-base bias and bias applied across the oxide. It was found that the SiO_2 transmission probability depends on the electron energy, and it was explained in terms of electron–phonon scattering in SiO_2. A good agreement with complementary Monte Carlo simulations that included details of the SiO_2 band structure and energy-dependent electron–phonon scattering effect was obtained [190].

In 30 Å Pd/75 Å SiO_2/Si(100), BEEM was applied to estimate the dielectric constant in SiO_2 layer by studying the barrier lowering (due to the image force effects) as a function of applied oxide bias [193].

6.1.2. Metal/CaF_2/Si MIS system

Cuberes et al. [194] have performed BEEM experiments in Au/14 Å CaF_2/Si(100) structure under UHV conditions. The authors observed a BEEM current onset at $V_t = 3.3$ V, and it was interpreted in terms of current transport through the CaF_2 conduction band. No irreversible loss of ballistic transmittance was observed (up to $V_t \sim 8$ V). Also the authors found a strong effect of the DOS of the CaF_2 film on the BEEM spectral shape, in reasonable accord with the theoretical predictions [195].

LaBella et al. [196,197] have characterized the hot electron transport properties of ultrathin Pt/CaF_2/Si(111) system by in-situ BEEM and STM. The 2–10 ML CaF_2 was grown epitaxially on Si(111) at 770°C, that gives many atomically flat, relatively defect-free terraces and steps. The Pt



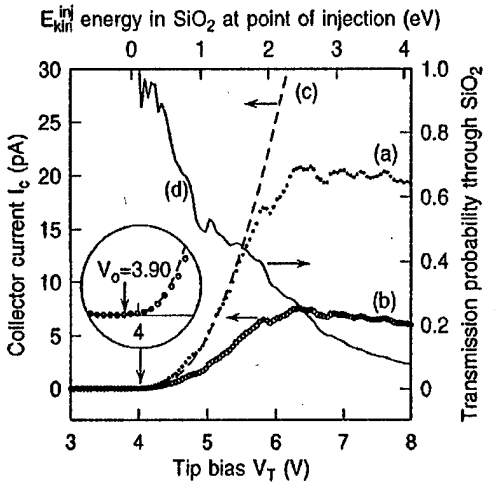


Fig. 52. (a) BEEM spectrum measured on a Au/Si sample without an oxide barrier, with $d_{Au} = 50$ Å and $I_t = 1$ nA. The solid line is a fit to the Bell and Kaiser model giving $eV_s = 0.76$ eV. (b) and (c) BEEM spectra measured on samples with a thin oxide barrier ($d_{ox} \approx 30$ Å). The solid lines are fits to a quadratic law $I_c \propto (V - V_s)^2$. (b) Case of a Au base with $d_{Au} = 100$ Å and $I_t = 10$ nA, giving $eV_s = 4.0$ eV. (c) Case of an Ir base with $d_{Ir} = 100$ Å and $I_t = 10$ nA, giving $eV_s = 5.6$ eV. The arrows indicate the threshold position [192].

Fig. 53. BEEM spectra of a 62 Å SiO_2 layer biased at $V_{ox} = 2$ V before [curve (a)] and after [curve (b)] the removal of contributions from impact ionization in the Si. Curve (c) is a model fit to the threshold region of (b) for an oxide transmission probability $T_{ox} = 1$. Curve (d), the ratio of (b)/(c), represents T_{ox} determined from the data. The inset depicts the fitted curve in the threshold region magnified 2× along the abscissa and 20× along the ordinate [190].

morphology consists of nodules, which nucleate at step edges and at defects ascribed to F vacancies on the CaF_2 surface. Changes in the tip bias provide a significant change in the imaged topography. The STM images at biases well above 3.3 eV (conduction band minimum of CaF_2) show similar features to bare CaF_2/Si(111), while images at the CBM show features of the deposited Pt. These topography results correlate well with the BEEM spectroscopy. Indeed, the BEEM spectra of 10 Å Pt/5 Å CaF_2/Si(111) show a peak at ~4.5 eV (the increase starts at ~3–3.5 eV) due to the DOS of the CaF_2 intralayer and an additional peak at 2 eV, indicating transmission through defects at the Pt/CaF_2 interface.

Recently, Sumya et al. [198,199] conducted BEEM studies of the electron transport phenomena across Au/CaF_2/n-Si(111) heterostructures as a function of the CaF_2 growth temperature (20°C, 550°C and 700°C). It was shown that the Au growth and electron transport properties depend strongly on the 2 ML CaF_2 growth temperature. The threshold voltage for an insulating CaF_2 intralayer, which is about 3.3–3.6 eV, is obtained only in the sample in which CaF_2 was deposited at 700°C, whereas for lower CaF_2 growth temperatures, the BEEM onset starts at ~1 and ~0.7 eV for 20°C and 550°C, respectively, as shown in Fig. 54. The complementary Fourier-transform infrared spectroscopy shows the presence of Ca–Si–F bonds for the 550°C deposition, and this implies that the CaF_2 heteroepitaxial growth at 550°C is unsuccessful in obtaining a high-quality CaF_2 layer. The nonepitaxial growth of CaF_2 at low temperatures creates defect states in the CaF_2 layer and/or at the interface below the CaF_2 conduction band that facilitate electron transport and lower the BEEM threshold.

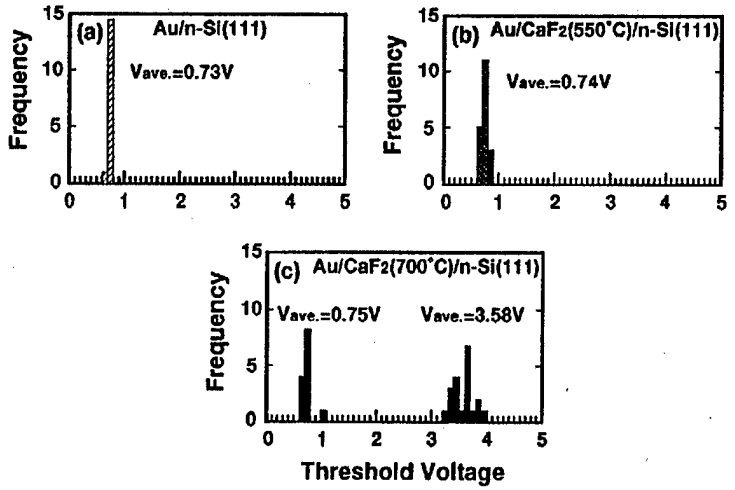


Fig. 54. Distributions of measured threshold voltages of (a) 25 ML Au/Si(1 1 1), (b) 50 ML Au/2 ML CaF_2 (550°C)/Si(111), and (c) 50 ML Au/2ML CaF_2 (700°C)/n-Si(1 1 1) [199].

6.2. *Charge trapping (build-up) and MOS device degradation*

A large portion of present reliability issues in MOS-based technology is related to the presence of hot electrons in the gate SiO_2 layers. Problems range from undesired buildup of trapping charge in the oxide film to degradation and breakdown of the dielectric.

BEEM was used to locally inject electrons into thick ~ 250 Å SiO_2 film built into a MOS structure [27,200]. Some of these electrons are trapped in the bulk of the oxide and cause an increase in the local barrier height and local time-dependent suppression of the BEEM current, consistent with a build-up of trapped BEEM electrons in the oxide. By studying the lateral extent of the current suppression caused by injection of charge at one point, the authors concluded that generation of traps created in the oxide by hot BEEM electrons is a likely dominant trapping mechanism. Measured variations in the BEEM threshold voltage with the voltage applied across the SiO_2 film can be used to estimate the local trapped charge density and the approximate depth of the trapped charge in the oxide.

The high current densities ($\sim 10^3$ C/cm^2 s) and choice of energy make BEEM an attractive method to study breakdown phenomena in dielectrics. By local stressing of thin SiO_2 layers (~ 25–40 Å) by injecting hot electrons into the conduction band of SiO_2, it is possible to inject very high charge densities of very hot electrons into the oxide film without damaging the MOS structure, as was found in $Pd/SiO_2/Si(100)$ [201] and $Pt/SiO_2/Si(100)$ [191]. A successful breakdown sequence for a 38 Å oxide stressed with 6 eV (kinetic energy) electrons is shown in Fig. 55. It was found that the total charge injected at breakdown exceeds by several orders of magnitude the values obtained by conventional Fowler–Nordheim tunnel injection under high field conditions. It was concluded that the observed breakdowns are still controlled by impurities/defects. Similar effect of the BEEM current build-up was also observed by Wen et al. [202] under even modest exposures (several single scans up to $V_t = 7$ V at $I_t = 2$ nA) in Pd/28 Å SiO_2/Si(100). This is in contrast to the suppression of the BEEM current observed on thicker SiO_2 films ($d > 40$ Å) due to

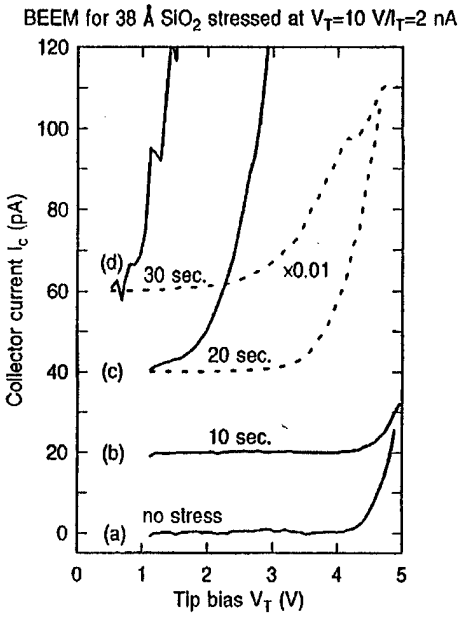


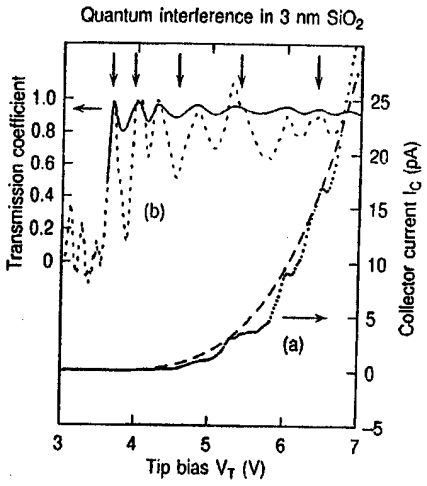
Fig. 55. A successful breakdown sequence for a 3.8 nm oxide stressed with 6 eV (kinetic energy) electrons. The shift to higher V_s between curve (a) and (b) is due to negative charge at the SiO_2–Si interface. Total injected charge to breakdown (defined by reaching $V_s \sim 1$ V, as in curve (c) is $\sim 6 \times 10^{13}$ C/cm^2 [201].

the buildup of negative charge in the oxide [27,200,203]. One possible explanation for this enhancement is related to an electron-leakage-induced buildup of positive charge at the SiO_2/Si interface accompanied by a decrease of the barrier in the oxide due to image force lowering effects. However, the experimental results are inconclusive because of relatively large noise in the BEEM current.

Degradation processes initiated by defect generation in device-grade SiO_2 were studied by locally injecting hot electrons from a STM tip into Pd/SiO_2/Si(100) MOS structure [204]. By analyzing the obtained BEEM results within the framework of a sheet charge model, charge densities in the oxide film were estimated. For the sample with 71 Å SiO_2 thickness, it was found that the SiO_2 film contains $\sim (0.7–2.8) \times 10^{13}$ cm^2 density of electron traps that are distributed within a 30 Å region adjacent to the Pd/SiO_2 interface. By additional stressing at high tip voltages up to 10 V, new traps characterized by a charge density of $\sim (1.9–3.6) \times 10^{13}$ cm^2 were created within a 40 Å region near the SiO_2/Si interface when the kinetic energy of the electron injected into the SiO_2 conduction band exceeds 1.9 eV. This energy is close to a minimum threshold of ~ 2 eV in order to break H–Si bonds. This result suggests that the initial defect generation, which eventually leads to a destructive breakdown of the oxide, is associated with the trap creation by releasing hydrogen at the anode.

6.3. *Quantum-size effect*

Kaczer et al. [191] have attempted to determine the SiO_2 layer thickness from the BEEM spectra. In the case of the BEEM transport above the SiO_2 CBM, the BEEM current is only a weak



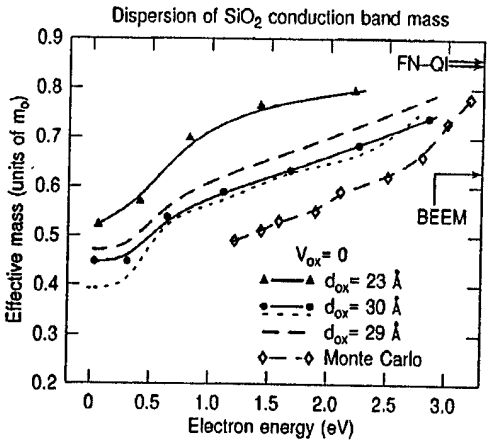


Fig. 56. (a) BEEM spectrum I_c V_t for 1.8 nm W/3.0 nm SiO_2/p-Si(1 0 0) MOS structure (dotted curve), with I_c^0 shown by a dashed line. (b) Experimental transmission coefficient, $t(E) = I_c/I_c^0$ (dotted curve), compared to $t(E)$ calculated with a dispersive mass (solid curve). Vertical arrows mark QI maxima for $t(E)$ calculated with a fixed mass $m_{ox} = 0.42\ m_0$. The threshold is at 3.77 V [205].

Fig. 57. Conduction band mass dispersions $m_{ox}(E)$ for SiO_2 determined from quantum interference oscillations in BEEM spectra (top four curves), compared to dispersion derived from Monte Carlo simulations of experimentally determined electron mean free paths (lowest dash-dotted curve) [205].

function of the film thickness. In the case of ultrathin SiO_2 (< 30 Å) when the tip voltage is below the SiO_2 CBM, a thickness-sensitive direct tunneling through the oxide will become considerable. However, in the BEEM experiment with 27 Å SiO_2, the BEEM current signal was too small to produce a measurable signal in this regime. It was discussed that the alternative method for estimating local thickness variations of the ultrathin SiO_2 layer is based on measurement of possible quantum-mechanical resonances of BEEM electrons in the conduction band of the SiO_2 film. The authors demonstrated that with some experimental care the oscillations could actually be observable in the first-voltage-derivative BEEM spectra (extracted numerically from the original BEEM spectra), in reasonable agreement with the theoretical expectations for hot electron resonance.

In 50 Å Pd/28 Å SiO_2/Si(100) system, quantum-interference oscillations of electrons (these oscillations arise from the constructive/deconstructive interference of electron waves reflected at the Pd/SiO_2 and SiO_2/Si interfaces of the SiO_2 cavity) in a thin SiO_2 layer were observed by BEEM using $I_t = 2$ nA [28,202]. Oscillations with up to four peaks in the energy region of 0–3 eV above the injection threshold were observed. The magnitude of the oscillations is $\sim 30\%$ of the underlying BEEM current. It was noted that the observation of pronounced oscillations due to the quantum-interference effects in SiO_2 is somewhat surprising because the SiO_2 film is quite leaky, with allowed states in both cladding regions (metal and Si) and extensive electron–phonon scattering losses in SiO_2 (electron mfp in SiO_2 at $T = 300$ K is ~ 10–20 Å).

Quantum interference oscillations in BEEM spectra were also observed for MOS structures with 23 and 30 Å SiO_2 interlayers [205,206]. The authors present an approach based on quantum interference effects of electrons injected directly into the conduction band of SiO_2 from which it is

possible to derive the energy dispersion of the electron effective mass m_{ox}. On the virgin surface, very well-pronounced oscillations (steps) in the BEEM current were observed during the first scan. However, with next scans at the same location, it was observed that the BEEM current increases and smoothens due to the stress-induced positive charging of the SiO_2 film at high voltages. For over-the-barrier transmission, maxima in the transmission probability for a rectangular barrier occur at the following energies $E = (n\pi\hbar/d_{ox})^2/2m_{ox}$ with $n = 1,2,3,\ldots$, from which m_{ox} can be deduced by matching theoretical maxima to those obtained experimentally (after correction on the SiO_2 band bending and image potential effects) (see Fig. 56). Using this approach, the 23 and 30 Å oxides exhibit initial (zero kinetic energy) m_{ox} values of $0.52m_0$ and $0.45m_0$, respectively, that disperse upwards by $\sim 0.3m_0$ over a 0–2.5 eV range in kinetic energies, as shown in Fig. 57.

7. Conclusion

In this article, we reviewed briefly recent progress in BEEM experiment and theory for the study of metal/semiconductor and metal/insulator/semiconductor devices. Excellent spatial resolution and independent control of the hot electron energy over a wide range make this technique a powerful tool for nanometer-scale characterization of the spatial and electronic properties of semiconductor (insulator) structures. In BEEM experiments, the hot electron transport is affected drastically by the integrated electron propagation and scattering in the metal layer, at the m–s interface and in the semiconductor itself. As a consequence, BEEM can be characterized by its spatial, energetic and depth resolutions.

Although the BEEM technique was originally invented as a unique microscopic and spectroscopic method to probe the Schottky barriers on a local scale, it has been successfully used for imaging and spectroscopy of buried quantum objects as well as for nondestructive local characterization of buried semiconductor heterostructures. Since BEEM spectra exhibit thresholds at energies where semiconductor states become available for transport, the SD-BEEM spectroscopy is very effective in identifying these thresholds and correlating them to semiconductor band structure. Significant progress was accomplished in establishing BEEM as an effective method for measuring semiconductor heterojunction offsets and for measuring resonant transport through double barrier and superlattice resonant tunneling heterostructures. Based on recent research, we demonstrated BEEM capability for imaging and electron spectroscopy of buried nanosized structures such as single self-assembled quantum dots as well as interfacial dislocations buried below the surface, by exploiting the lateral resolution of BEEM.

Acknowledgements

We would like to thank our collaborators S. Bhargava, E.G. Brazel, M.A. Chin, E.Y. Lee, J.J. O'Shea, C.V. Reddy, M.E. Rubin, T. Sajoto and C. Zheng, with whom some original parts of the work reported here were done. We also thank our collaborators D.L. Smith of Los Alamos National Laboratory, A. Mascarenhas, Y. Zhang and J.M. Olson of National Renewable Energy Laboratory, H.P. Xin and C.W. Tu of University of California at San Diego, Yi-Jen Chiu and J.E. Bowers of University of California at Santa Barbara and R.D. Dupuis of University of Texas at

Austin for helpful discussions. This work was supported by NSF under Grants # ECS 99-96093, ECS 99-06047 and # DMR 98-09363, AFOSR under Grant # F49620-97-10247. We also thank the Midwest Research Institute, the U.S. Army Research Office under Award # DAAD 19-99-1-0158 and the Regents of the University of California for partial support.

References

[1] L.D. Bell, W.J. Kaiser, Phys. Rev. Lett. 61 (1988) 2368 ***.
[2] M. Prietsch, Phys. Rep. 253 (1995) 163 ***.
[3] L.D. Bell, W.J. Kaiser, Annu. Rev. Mater. Sci. 26 (1996) 189 **.
[4] Fernandez, H.D. Hallen, T. Huang, R.A. Buhrman, J. Silcox, Appl. Phys. Lett. 57 (1990) 2826 **.
[5] M. Kozhevnikov, V. Narayanamurti, C. Zheng, Y.-J. Chiu, D.L. Smith, Phys. Rev. Lett. 82 (1999) 3677 **.
[6] P. Niedermann, L. Quattropani, K. Solt, A.D. Kent, O. Fischer, J. Vac. Sci. Technol. B 10 (1992) 580.
[7] Bauer, M.T. Cuberes, M. Prietsch, G. Kaindl, Phys. Rev. Lett. 71 (1993) 149.
[8] H. Sirringhaus, E.Y. Lee, H. von Kanel, Surf. Sci. 331–333 (1995) 1277.
[9] K. Reuter, F.J. Garcia-Vidal, P.L. de Andres, F. Flores, K. Heinz, Phys. Rev. Lett. 81 (1998) 4963 *.
[10] R.H. Williams, Appl. Surf. Sci. 70–71 (1993) 386.
[11] L.D. Bell, S.J. Manion, M.H. Hecht, W.J. Kaiser, R.W. Fathauer, A.M. Milliken, Phys. Rev. B 48 (1993) 5712.
[12] L.D. Bell, A.M. Milliken, S.J. Manion, W.J. Kaiser, R.W. Fathauer, W.T. Pike, Phys. Rev. B 50 (1994) 8021.
[13] R. Ludeke, Phys. Rev. Lett. 70 (1993) 214.
[14] Chahboun, R. Coratger, F. Ajustron, J. Beauvillain, I.M. Dharmadasa, A.P. Samantilleke, J. Appl. Phys. 87 (2000) 2422.
[15] J.J. O'Shea, C.M. Reaves, S.P. DenBaars, M.A. Chin, V. Narayanamurti, Appl. Phys. Lett. 69 (1996) 3022.
[16] M. Kozhevnikov, V. Narayanamurti, A. Mascarenhas, Y. Zhang, J.M. Olson, D.L. Smith, Appl. Phys. Lett. 75 (1999) 1128.
[17] L.D. Bell, R.P. Smith, B.T. McDermott, E.R. Gertner, R. Pittman, R.L. Pierson, G.L. Sullivan, J. Vac. Sci. Technol. B 16 (1998) 2286.
[18] E. Brazel, M.A. Chin, V. Narayanamurti, Appl. Phys. Lett. 74 (1999) 2367.
[19] L.D. Bell, R.P. Smith, B.T. McDermott, E.R. Gertner, R. Pittman, R.L. Pierson, G.J. Sullivan, Appl. Phys. Lett. 76 (2000) 1725.
[20] M. Kozhevnikov, V. Narayanamurti, C.V. Reddy, H.P. Xin, C.W. Tu, A. Mascarenhas, Y. Zhang, Phys. Rev. B 61 (2000) R7861.
[21] H.-J. Im, B. Kaczer, J.P. Pelz, W.J. Choyke, Appl. Phys. Lett. 72 (1998) 839.
[22] B. Kaczer, H.-J. Im, J.P. Pelz, J. Chen, W.J. Choyke, Phys. Rev. B 57 (1998) 4027.
[23] W.J. Kaiser, L.D. Bell, Phys. Rev. Lett. 60 (1988) 1406.
[24] For review, see, e.g., V. Narayanamurti, Sci. Rep. Res. Inst. Tokohu Univ. A, Phys. Chem. Metall. 44 (1997) 165.
[25] M.T. Cuberes, A. Bauer, H.J. Wen, M. Prietsch, G. Kaindl, J. Vac. Sci. Technol. B 12 (1994) 2646.
[26] R. Ludeke, A. Bauer, E. Cartier, Appl. Phys. Lett. 66 (1995) 730.
[27] B. Kaczer, J.P. Pelz, J. Vac. Sci. Technol. B 14 (1996) 2864.
[28] R. Ludeke, H.J. Wen, A. Schenk, Appl. Phys. Lett. 73 (1998) 1221.
[29] J.J. O'Shea, E.G. Brazel, M.E. Rubin, S. Bhargava, M.A. Chin, V. Narayanamurti, Phys. Rev. B 56 (1997) 2026 **.
[30] S. Bhargava, C. Zheng, J. Ko, M.A. Chin, L.A. Coldren, V. Narayanamurti, Appl. Phys. Lett. 73 (1998) 3271.
[31] S. Bhargava, H.-R. Blank, E. Hall, M.A. Chin, H. Kroemer, V. Narayanamurti, Appl. Phys. Lett. 74 (1999) 1135.
[32] D.K. Guthrie, P.N. First, T.K. Gaylord, E.N. Glytsis, R.E. Leibenguth, Appl. Phys. Lett. 71 (1997) 2292.
[33] T. Sajoto, J.J. O'Shea, S. Bhargava, D. Leonard, M.A. Chin, V. Narayanamurti, Phys. Rev. Lett. 74 (1995) 3427 **.
[34] R. Heer, J. Smoliner, G. Strasser, E. Gornik, Appl. Phys. Lett. 73 (1998) 3138.
[35] J. Smoliner, R. Heer, C. Eder, G. Strasser, Phys. Rev. B 58 (1998) R7516.
[36] J. Smoliner, C. Eder, G. Böhm, G. Weimann, Appl. Phys. Lett. 69 (1996) 52.
[37] C. Eder, J. Smoliner, G. Strasser, Appl. Phys. Lett. 68 (1996) 2876.

[38] M.E. Rubin, G. Medeiros-Ribeiro, J.J. O'Shea, M.A. Chin, E.Y. Lee, P.M. Petroff, V. Narayanamurti, Phys. Rev. Lett. 77 (1996) 5268 **.
[39] M.E. Rubin, H.R. Blank, M.A. Chin, H. Kroemer, V. Narayanamurti, Appl. Phys. Lett. 70 (1997) 1590.
[40] C.V. Reddy, V. Narayanamurti, J.H. Ryou, U. Chowdhury, R.D. Dupuis, Appl. Phys. Lett. 76 (2000) 1437.
[41] H. Sirringhaus, E.Y. Lee, U. Kafader, H. von Kanel, J. Vac. Sci. Technol. B 13 (1995) 1848.
[42] E.Y. Lee, S. Bhargava, K. Pond, K. Luo, M.A. Chin, V. Narayanamurti, Appl. Phys. Lett. 69 (1996) 940.
[43] E.Y. Lee, S. Bhargava, M.A. Chin, V. Narayanamurti, J. Vac. Sci. Technol. A 15 (1997) 1351 **.
[44] A. Bauer, R. Ludeke, J. Vac. Sci. Technol. B 12 (1994) 2667.
[45] G.N. Henderson, P.N. First, T.K. Gaylord, E.N. Glytsis, B.J. Rice, P.L. Dantzscher, D.K. Guthrie, L.E. Harrell, J.S. Cave, Rev. Sci. Instr. 66 (1995) 91.
[46] C. Eder, J. Smoliner, G. Strasser, E. Gornik, Appl. Phys. Lett. 69 (1996) 1725.
[47] J.G. Simmons, J. Appl. Phys. 34 (1963) 1793.
[48] R. Ludeke, M. Prietsch, J. Vac. Sci. Technol. A 9 (1991) 885 **.
[49] Mao-long Ke, D.I. Westwood, C.C. Matthai, B.E. Richardson, R.H. Williams, Phys. Rev. B 53 (1996) 4845.
[50] G.N. Henderson, P.N. First, T.K. Gaylord, E.N. Glytsis, Phys. Rev. Lett. 71 (1993) 2999.
[51] M.T. Cuberes, A. Bauer, H.J. Wen, D. Vandre, M. Prietsch, G. Kaindl, J. Vac. Sci. Technol. B 12 (1994) 2422.
[52] E.Y. Lee, L.J. Schowalter, J. Appl. Phys. 66 (1989) 4903.
[53] R. Menegozzi, P.-G. Reinhard, M. Schulz, Appl. Phys. A 66 (1998) S897 **.
[54] R. Menegozzi, P.-G. Reinhard, M. Schulz, Surf. Sci. 411 (1998) L810.
[55] C.A. Ventrice, V.P. LaBella, G. Ramaswamy, H.-P. Yu, L.J. Schowalter, Phys. Rev. B 53 (1996) 3952.
[56] L.J. Schowalter, E.Y. Lee, Phys. Rev. B 43 (1991) 9308.
[57] C. Manke, Y. Bodschwinna, M. Schulz, Appl. Surf. Sci. 117/118 (1997) 321.
[58] D.L. Smith, S.M. Kogan, Phys. Rev. B 54 (1996) 10354 ***.
[59] D.L. Smith, M. Kozhevnikov, E.Y. Lee, V. Narayanamurti, Phys. Rev. B 61 2000 13914 **.
[60] M.H. Hecht, L.D. Bell, W.J. Kaiser, L.C. Davis, Phys. Rev. B 42 (1990) 7663.
[61] L.D. Bell, W.J. Kaiser, M.H. Hecht, L.C. Davis, in: J.A. Stroscio, W.J. Kaiser (Eds.), Methods of Experimental Physics 27: Scanning Tunneling Microscopy, Academic Press, New York, 1993.
[62] E.Y. Lee, L.J. Schowalter, Phys. Rev. B 45 (1992) 6325.
[63] E.Y. Lee, H. Sirringhaus, U. Kafader, H. von Kanel, Phys. Rev. B 52 (1995) 1816.
[64] R. Ludeke, A. Bauer, J. Vac. Sci. Technol. A 13 (1995) 614 ***.
[65] E. Cartier, M.V. Fischetti, E.A. Eklund, F.R. MsFeely, Appl. Phys. Lett. 62 (1993) 3339.
[66] R.C. Alig, S. Bloom, C.W. Struck, Phys. Rev. B 22 (1980) 5565.
[67] E.Y. Lee, L.J. Schowalter, J. Appl. Phys. 70 (1991) 2156 **.
[68] A. Davies, H.G. Craighead, Appl. Phys. Lett. 64 (1994) 2833.
[69] K. Reuter, U. Hohenester, P.L. de Andres, F.J. Garcia-Vidal, F. Flores, K. Heinz, P. Kocevar, Phys. Rev. B 61 (2000) 4522.
[70] L.J. Schowalter, E.Y. Lee, Phys. Rev. B 43 (1991) 9308.
[71] H.D. Hallen, A. Fernandez, T. Huang, J. Silcox, R.A. Buhrman, Phys. Rev. B 46 (1992) 7256.
[72] A. Fernandez, H.D. Hallen, T. Huang, R.A. Buhrman, J. Silcox, Phys. Rev. B 44 (1991) 3428.
[73] A. Fernandez, H.D. Hallen, T. Huang, R.A. Buhrman, J. Silcox, J. Vac. Sci. Technol. B 9 (1991) 590.
[74] W.L. Kaiser, M.H. Hecht, R.W. Fathauer, L.D. Bell, E.Y. Lee, L.C. Davis, Phys. Rev. B 44 (1991) 6546.
[75] H. Sirringhaus, E.Y. Lee, H. von Kanel, J. Vac. Sci. Technol. B 12 (1994) 2629.
[76] M.D. Stiles, D.R. Harmann, J. Vac. Sci. Technol. B 9 (1991) 2394.
[77] T. Meyer, D. Migas, L. Miglio, H. von Kanel, Phys. Rev. Lett. 85 (2000) 1520; H. Sirringhaus, E.Y. Lee, H. von Kanel, Phys. Rev. Lett. 73 (1994) 577 ***.
[78] T. Meyer, H. Sirringhaus, H. von Kanel, Thin Solid Films 318 (1998) 195.
[79] E.Y. Lee, H. Sirringhaus, H. von Kanel, Phys. Rev. B 50 (1994) 5807.
[80] H. von Kanel, E.Y. Lee, H. Sirringhaus, U. Kafader, Thin Solid Films 267 (1995) 89.
[81] H. Sirringhaus, T. Meyer, E.Y. Lee, H. von Kanel, Phys. Rev. B 53 (1996) 15944 **.
[82] H. von Kanel, T. Meyer, H. Sirringhaus, Jpn. J. Appl. Phys. 37 (1998) 3800 **.
[83] M.D. Stiles, D.R. Hamann, Phys. Rev. Lett. 66 (1991) 3179.

[84] D.K. Guthrie, L.E. Harrell, G.N. Henderson, P.N. First, T.K. Gaylord, E.N. Glytsis, R.E. Leibenguth, Phys. Rev. B 54 (1996) 16972.
[85] D.L. Smith, E.Y. Lee, V. Narayanamurti, Phys. Rev. Lett. 80 (1998) 2433 ***.
[86] L.D. Bell, Phys. Rev. Lett. 77 (1996) 3893 *.
[87] L.D. Bell, J. Vac. Sci. Technol. A 15 (1997) 1358.
[88] P.M. Solomon, S.L. Wright, C. Lanza, Superlattices Microstruct. 2 (1986) 521.
[89] E.E. Mendez, E. Calleja, C.E.T.G. da Silva, L.L. Ghang, W.I. Wang, Phys. Rev. B 33 (1986) 7368.
[90] H.C. Liu, Appl. Phys. Lett. 51 (1987) 1019.
[91] Z.S. Gribnikov, O.E. Raichev, Soc. Phys. Semicond. 23 (1989) 1344.
[92] C. Girardin, R. Coratger, R. Pechou, F. Ajustron, J. Beauvillain, J. Phys. III France 6 (1996) 661.
[93] M.K. Weilmeier, W.H. Rippard, R.A. Buhrman, Phys. Rev. B 59 (1999) R2521 **.
[94] F.J. Garcia-Vidal, P.L. de Andres, F. Flores, Phys. Rev. Lett. 76 (1996) 807.
[95] K. Reuter, P.L. de Andres, F.J. Garcia-Vidal, D. Sestovic, F. Flores, K. Heinz, Phys. Rev. B 58 (1998) 14036 *.
[96] L.D. Bell, M.H. Hecht, W.J. Kaiser, L.C. Davis, Phys. Rev. Lett. 64 (1990) 2679.
[97] W.H. Rippard, R.A. Buhrman, Phys. Rev. Lett. 84 (2000) 971.
[98] H. Sirringhaus, E.Y. Lee, H. von Kanel, Phys. Rev. Lett. 74 (1995) 3999 ***.
[99] A. Davies, J.G. Couillard, H.G. Craighead, Appl. Phys. Lett. 61 (1992) 1040.
[100] A. Olbrich, J. Vancea, F. Kreupl, H. Hoffmann, J. Appl. Phys. 83 (1998) 358.
[101] E.Y. Lee, H. Sirringhaus, H. von Kanel, Surf. Sci. Lett. 314 (1994) L823.
[102] A.M. Milliken, S.J. Manion, W.J. Kaiser, L.D. Bell, M.H. Hecht, Phys. Rev. B 46 (1992) 12826.
[103] U. Hohenester, P. Kacevar, P. de Andres, F. Flores, Phys. Stat. Sol. (b) 204 (1997) 397.
[104] B.A. Morgan, K.M. Ring, K.L. Kavanagh, A.A. Talin, R.S. Williams, T. Yasuda, T. Yasui, Y. Segawa, J. Vac. Sci. Technol. B 14 (1996) 1238.
[105] J.G. Nunes, N.M. Amer, Appl. Phys. Lett. 63 (1993) 1851.
[106] M.W. Prins, R. Jansen, H. van Kempen, Phys. Rev. B 53 (1996) 8105.
[107] L.O. Olsson, C.B.M. Andersson, M.C. Hakansson, J. Kanski, L. Ilver, O.U. Karlsson, Phys. Rev. Lett. 76 (1996) 3626.
[108] J. Smoliner, C. Eder, Phys. Rev. B 57 (1998) 9856.
[109] J. Smoliner, R. Heer, C. Eder, Appl. Phys. A 66 (1998) S117.
[110] E.Y. Lee, Phys. Rev. B 59 (1999) 15332.
[111] E.Y. Lee, V. Narayanamurti, D.L. Smith, Phys. Rev. B 55 (1997) R16033.
[112] S. Bhargava, Ph.D. Dissertation, University of California at Santa Barbara, August 1998.
[113] M.H. Hecht, L.D. Bell, W.J. Kaiser, F.J. Grunthaner, Appl. Phys. Lett. 55 (1989) 780.
[114] G.M. Vanalme, L. Goubert, R.L.V. Meirhaeghe, F. Cardon, P.V. Daele, Semicond. Sci. Technol. 14 (1999) 871.
[115] C. Detavernier, R.L.V. Meirhaeghe, R. Donaton, K. Maex, F. Cardon, J. Appl. Phys. 84 (1998) 3226.
[116] J.L. Everaert, R.L.V. Meirhaeghe, W.H. Laflere, F. Cardon, Semicond. Sci. Technol. 10 (1995) 504.
[117] G.M. Vanalme, R.L.V. Meirhaeghe, F. Cardon, P.V. Daele, Semicond. Sci. Technol. 12 (1997) 907.
[118] L. Stockman, H. van Kempen, Surf. Sci. 408 (1998) 232.
[119] I.M. Dharmadasa, C.J. Blomfield, C.G. Scott, R. Coratger, F. Ajustron, J. Beauvillain, Solid State Electron. 42 (1998) 595.
[120] H. Sirringhaus, T. Meyer, E.Y. Lee, H. von Kanel, Surf. Sci. 357–358 (1996) 386.
[121] Y. Hasegawa, K. Akiyama, M. Ono, S.-J. Kahng, Q.K. Xue, K. Nakayama, T. Hashizume, T. Sakurai, Appl. Phys. Lett. 75 (1999) 3668 *.
[122] M.K. Weilmeier, W.H. Rippard, R.A. Buhrman, Phys. Rev. B 61 (2000) 7161.
[123] M. Prietsch, R. Ludeke, Phys. Rev. Lett. 66 (1991) 2511.
[124] H.D. Hallen, A. Fernandez, T. Huang, R.A. Buhrman, J. Silcox, J. Vac. Sci. Technol. B 9 (1991) 585.
[125] X. Qiu, G. Shang, C. Wang, C. Bai, Appl. Phys. A 66 (1998) S91.
[126] G.N. Henderson, T.K. Gaylord, E.N. Glytsis, P.N. First, W.J. Kaiser, Solid State Commun. 80 (1991) 591.
[127] R. Coratger, C. Girardin, J. Beauvillain, I.M. Dharmadasa, A.P. Samanthilake, J.E.F. Frost, J. Appl. Phys. 81 (1997) 7870.

[128] B.A. Morgan, K.M. Ring, K.L. Kavanagh, A.A. Talin, R.S. Williams, T. Yasuda, T. Yasui, Y. Segawa, J. Appl. Phys. 79 (1996) 1532.
[129] R. Coratger, F. Ajustron, J. Beauvillain, I.M. Dharmadasa, C.J. Blomfeld, K.A. Prior, J. Simpson, B.C. Cavenett, Phys. Rev. B 51 (1995) 2357.
[130] L.D. Bell, W.J. Kaiser, S.J. Manion, A.M. Milliken, R.W. Fathauer, W.T. Pike, Phys. Rev. B 52 (1995) 12,081.
[131] C.G. Van de Walle, R.M. Martin, Phys. Rev. B 34 (1986) 5621.
[132] Mao-long Ke, D.I. Westwood, C.C. Matthai, R.H. Williams, Surf. Sci. 352–354 (1996) 861.
[133] Mao-long Ke, D.I. Westwood, C.C. Matthai, B.E. Richardson, R.H. Williams, J. Vac. Sci. Technol. B 14 (1996) 2786.
[134] S. Bhargava, H.-R. Blank, V. Narayanamurti, H. Kromer, Appl. Phys. Lett. 70 (1997) 759.
[135] X.-C. Cheng, T.C. McGill, J. Vac. Sci. Technol. B 16 (1998) 2291.
[136] S.D. Lester, F.A. Ponce, M.G. Craford, D.A. Steigerwald, Appl. Phys. Lett. 66 (1995) 1249.
[137] L.D. Bell, R.P. Smith, B.T. McDermott, E.R. Gertner, R. Pittman, R.L. Pierson, G.L. Sullivan, Appl. Phys. Lett. 72 (1998) 1590.
[138] S.-H. Wei, A. Zunger, Appl. Phys. Lett. 56 (1990) 662.
[139] S.-H. Wei, D.B. Laks, A. Zunger, Appl. Phys. Lett. 62 (1993) 1937.
[140] Y. Zhang, A. Mascarenhas, P. Ernst, F.A.J.M. Driessen, D.J. Friedman, C. Geng, F. Scholz, H. Schweizer, J. Appl. Phys. 81 (1997) 6365.
[141] M.C. DeLong, D.J. Mowbray, R.A. Hogg, M.S. Skolnick, J.E. Williams, K. Meehan, S.R. Kurtz, J.M. Olson, R.P. Schneider, M.C. Wu, M. Hopkinson, Appl. Phys. Lett. 66 (1995) 3185.
[142] P. Ernst, C. Geng, F. Scholz, H. Schweizer, Y. Zhang, A. Mascarenhas, Appl. Phys. Lett. 67 (1995) 2347.
[143] B.T. McDermott, K.G. Reid, N.A. El-Masry, S.M. Bedair, W.M. Duncan, X. Yin, F.H. Pollak, Appl. Phys. Lett. 56 (1990) 1172.
[144] G.B. Stringfellow, G.S. Chen, J. Vac. Sci. Technol. B 9 (1991) 2182.
[145] D.J. Friedman, J.G. Zhu, A.E. Kibber, J.M. Olson, J. Moreland, Appl. Phys. Lett. 63 (1993) 1774.
[146] S.R. Kurtz, J. Appl. Phys. 74 (1993) 4130.
[147] S.H. Kwok, P.Y. Yu, K. Uchida, Phys. Rev. B 58 (1998) R13395.
[148] T. Kippenberg, J. Krauss, J. Spieler, P. Kiesel, G.H. Dohler, R. Stubner, R. Winkler, O. Pankratov, M. Moser, Phys. Rev. B 60 (1999) 4446.
[149] A. Zunger, private communications.
[150] S.-H. Wei, A. Franceschetti, A. Zunger, Phys. Rev. B 51 (1995) 13,097.
[151] S. Sakai, Y. Ueta, Y. Terauchi, Jpn. J. Appl. Phys. 32 (1993) 4413.
[152] M. Kondow, K. Uomi, K. Hosomi, T. Mozume, Jpn. J. Appl. Phys. 33 (1994) L1056.
[153] S.-H. Wei, A. Zunger, Phys. Rev. Lett. 76 (1996) 664.
[154] J. Salzman, H. Temkin, Mater. Sci. Eng. B 50 (1997) 148.
[155] W.G. Bi, C.W. Tu, Appl. Phys. Lett. 70 (1997) 1608.
[156] M. Kondow, K. Uomi, A. Niwa, T. Kitatani, S. Watahiki, Y. Yazawa, Jpn. J. Appl. Phys. 35 (1996) 1237.
[157] D.J. Friedman, J.F. Geisz, S.R. Kurtz, J.M. Olson, J. Crystal Growth 195 (1998) 409.
[158] H.P. Xin, C.W. Tu, Appl. Phys. Lett. 72 (1998) 2442.
[159] H. Gruning, L. Chen, T. Hartmann, P.J. Klar, W. Heimbrodt, F. Hohnsdorf, J. Koch, W. Stolz, Phys. Stat. Sol. (b) 215 (1999) 39.
[160] S. Francoeur, S.A. Nikishin, C. Jin, Y. Qui, H. Temkin, Appl. Phys. Lett. 75 (1999) 1538.
[161] W. Shan, W. Walukiewicz, J.W. Ager III, E.E. Haller, J.F. Geisz, D.J. Friedman, J.M. Olson, S.R. Kurtz, J. Appl. Phys. 86 (1999) 2349.
[162] J.D. Perkins, A. Mascarenhas, Y. Zhang, J.F. Geisz, D.J. Friedman, J.M. Olson, S.R. Kurtz, Phys. Rev. Lett. 82 (1999) 3312.
[163] T. Mattila, S.-H. Wei, A. Zunger, Phys. Rev. B 60 (1999) R11,245.
[164] N.A. Modine, unpublished.
[165] W. Shan, W. Walukiewicz, J.W. Ager III, E.E. Haller, J.F. Geisz, D.J. Friedman, J.M. Olson, S.R. Kurtz, Phys. Rev. Lett. 82 (1999) 1221.
[166] K. Uesugi, I. Suemune, Jpn. J. Appl. Phys. 36 (1997) L1572.

[167] G. Pozina, I. Ivanova, B. Monemar, J. Thordson, T.G. Andersson, Mater. Sci. Eng. B 50 (1997) 153.
[168] X.-C. Cheng, D.A. Collins, T.C. McGill, J. Vac. Sci. Technol. A 15 (1997) 2063.
[169] J.J. O'Shea, C.M. Reaves, M.A. Chin, S.P. DenBaars, A.C. Gossard, V. Narayanamurti, E.D. Jones, Mat. Res. Soc. Symp. Proc. 417 (1996) 79.
[170] S.L. Feng, J. Krynicki, V. Donchev, J.C. Bourgoin, M. Difortepoisson, C. Brylinski, S. Delage, H. Blanck, S. Alaya, Semicond. Sci. Technol. 8 (1993) 2092.
[171] D.K. Guthrie, P.N. First, T.K. Gaylord, E.N. Glytsis, R.E. Leibenguth, Appl. Phys. Lett. 72 (1998) 374 ***.
[172] D.K. Guthrie, P.N. First, T.K. Gaylord, E.N. Glytsis, R.E. Leibenguth, Appl. Phys. Lett. 75 (1999) 283.
[173] C. Eder, J. Smolimer, R. Heer, G. Strasser, E. Gornic, Physica E 2 (1998) 850.
[174] J. Smoliner, R. Heer, G. Strasser, Phys. Rev. B 60 (1999) R5137 *.
[175] C. Rauch, G. Strasser, K. Unterrainer, W. Boxleitner, E. Gornik, Phys. Stat. Sol. (b) 204 (1997) 393.
[176] H. von Kanel, T. Meyer, H. Sirringhaus, J. Crystal Growth 175–176 (1997) 340.
[177] J. Smoliner, C. Eder, G. Strasser, E. Gornic, Phys. Stat. Sol. (b) 204 (1997) 386.
[178] D.I. Westwood, Mao-long Ke, F. LeLarge, F. Laruelle, B. Etienne, Surf. Sci. 352–354 (1996) 802.
[179] F. Hatami, N.N. Ledentsov, M. Grundmann, J. Bohrer, F. Heinrichsdorff, M. Beer, D. Bimberg, S.S. Ruvimov, P. Werner, U. Gosele, J. Heydenreich, U. Richter, S.V. Ivanov, B.Ya. Meltser, P.S. Kop'ev, Zh.I. Alferov, Appl. Phys. Lett. 67 (1995) 656.
[180] B.R. Bennett, R. Magno, B.V. Shanabrook, Appl. Phys. Lett. 68 (1996) 505.
[181] C.-K. Sun, G. Wang, J.E. Bowers, B. Brar, H.-R. Blank, H. Kroemer, M.H. Pilkuhn, Appl. Phys. Lett. 68 (1996) 1543.
[182] C.V. Reddy, V. Narayanamurti, J.H. Ryou, U. Chowdhury, R.D. Dupuis, Appl. Phys. Lett. 77 (2000) 1167.
[183] T. Kinno, K. Tanaka, K. Mizushima, Phys. Rev. B 56 (1997) R4391.
[184] W.H. Rippard, R.A. Buhrman, Appl. Phys. Lett. 75 (1999) 1001 **.
[185] S.M. Sze, Physics of Semiconductor Devices, 2nd Edition, Wiley, New York, 1981.
[186] D. Landheer, Y. Tao, D.X. Xu, G.I. Sproule, D.A. Buchanan, J. Appl. Phys. 78 (1995) 1818.
[187] E. Cartier, J.C. Tsang, M.V. Fischetti, D.A. Buchanan, Microelectron. Eng. 36 (1997) 103.
[188] R.W. Keyes, The Physics of VLSI Systems, Addison-Wesley, New York, 1987.
[189] M. Dahle-Prietsch, I. Manke, T. Kalka, H.J. Wen, G. Kaindl, J. Phys. D: Appl. Phys. 30 (1997) L48.
[190] R. Ludeke, A. Bauer, E. Cartier, J. Vac. Sci. Technol. B 13 (1995) 1830 **.
[191] B. Kaczer, H.-J. Im, J.P. Pelz, J. Vac. Sci. Technol. B 16 (1998) 2302.
[192] L. Quattropani, I. Maggio-Aprile, P. Niedermann, O. Fisher, Phys. Rev. B 57 (1998) 6623.
[193] H.J. Wen, R. Ludeke, D.M. Newns, S.H. Lo, J. Vac. Sci. Technol. A 15 (1997) 784.
[194] M.T. Cuberes, A. Bauer, H.J. Wen, M. Prietsch, G. Kaindl, Appl. Phys. Lett. 64 (1994) 2300.
[195] C. Arcangeli, S. Ossicini, O. Bisi, Surf. Sci. 269–270 (1992) 743.
[196] V.P. LaBella, L.J. Schowalter, J.C.A. Ventrice, J. Vac. Sci. Technol. B 15 (1997) 1191.
[197] V.P. LaBella, Y. Shusterman, L.J. Schowalter, J.C.A. Ventrice, J. Vac. Sci. Technol. A 16 (1998) 1692.
[198] T. Sumiya, T. Miura, S.-I. Tanaka, J. Vac. Sci. Technol. A 16 (1998) 2653.
[199] T. Sumiya, K. Honda, T. Miura, S.-I. Tanaka, J. Appl. Phys. 85 (1999) 941.
[200] B. Kaczer, Z. Meng, J.P. Pelz, Phys. Rev. Lett. 77 (1996) 91.
[201] R. Ludeke, H.J. Wen, E. Cartier, J. Vac. Sci. Technol. B 14 (1996) 2855 *.
[202] H.J. Wen, R. Ludeke, A. Schenk, J. Vac. Sci. Technol. B 16 (1998) 2296.
[203] B. Kaczer, H.-J. Im, J.P. Pelz, R.M. Wallace, Appl. Phys. Lett. 73 (1998) 1871.
[204] H.J. Wen, R. Ludeke, J. Vac. Technol. B 15 (1997) 1080.
[205] R. Ludeke, E. Cartier, A. Schenk, Appl. Phys. Lett. 75 (1999) 1407 *.
[206] R. Ludeke, A. Schenk, J. Vac. Sci. Technol. B 17 (1999) 1823.

Design and characterization of thin film microcoolers

Chris LaBounty[a)]
Department of Electrical and Computer Engineering, University of California, Santa Barbara, California 93106

Ali Shakouri
Jack Baskin School of Engineering, University of California, Santa Cruz, California 95046

John E. Bowers
Department of Electrical and Computer Engineering, University of California, Santa Barbara, California 93106

(Received 29 February 2000; accepted for publication 15 January 2001)

Thin film coolers can provide large cooling power densities compared to bulk thermoelectrics due to the close spacing of hot and cold junctions. Important parameters in the design of such coolers are investigated theoretically and experimentally. A three-dimensional (3D) finite element simulator (ANSYS) is used to model self-consistently thermal and electrical properties of a complete device structure. The dominant three-dimensional thermal and electrical spreading resistances acquired from the 3D simulation are also used in a one-dimensional model (MATLAB) to obtain faster, less rigorous results. Heat conduction, Joule heating, thermoelectric and thermionic cooling are included in these models as well as nonideal effects such as contact resistance, finite thermal resistance of the substrate and the heat sink, and heat generation in the wire bonds. Simulations exhibit good agreement with experimental results from InGaAsP-based thin film thermionic emission coolers which have demonstrated maximum cooling of 1.15 °C at room temperature. With the nonideal effects minimized, simulations predict that single stage thin film coolers can provide up to 20–30 °C degrees centigrade cooling with cooling power densities of several 1000 W/cm^2. © *2001 American Institute of Physics.* [DOI: 10.1063/1.1353810]

I. INTRODUCTION

Temperature control of microelectronic and optoelectronic components is typically accomplished with thermoelectric (TE) coolers. TE coolers have become essential in modern optical telecommunications to control the characteristics of laser sources, switching/routing elements, and detectors used in wavelength division multiplexed systems. Cooling requirements in microprocessors and other integrated circuits have also risen dramatically in recent years due to the increase in clock speed and reduction in feature size. Generally, as these devices have become smaller, faster, and more dense, the power density has greatly increased. Conventional TE coolers are incompatible with integrated circuit fabrication processes, and are therefore limited in how small they can be manufactured. This bulk fabrication technology makes integration with microelectronic and optoelectronic devices difficult,[1] resulting in a high cost of packaging. Furthermore, the reliability of packaged modules employing a TE cooler is usually limited by the reliability of the cooler itself.[2] A solution to these problems is to shift from bulk thermoelements to integrated thin film coolers.

The greatest advantage of thin film coolers is the dramatic gain in cooling power density as it is inversely proportional to the length of the thermoelements. Thin films on the order of microns should provide cooling power densities greater than 1000 W/cm^2. In addition, several methods such as thermionic emission in heterostructures[3,4] and decreased thermal conductivity in superlattices,[5,6] are being explored to improve the performance beyond what is possible with bulk thermoelements. Thin film coolers can also be made in large quantities using well known integrated circuit batch fabrication methods resulting in lowered cost and greater reliability. Integration with microelectronic and optoelectronic devices is also possible for active localized cooling.

Many nonidealities become apparent and must be considered when moving from bulk to thin film coolers. Whereas contact resistance, thermal resistance of the heat sink, and heat generation in the current carrying connections are secondary effects in bulk TE coolers, they can all become critical in thin film coolers. In the following, these nonideal effects, as well as heat conduction, Joule heating, thermoelectric and thermionic cooling, are investigated. A three-dimensional (3D) self-consistent thermal/electrical software simulator[7] is used to model these effects (Fig. 1), and the results are compared to experimental measurements. The three-dimensional analysis is necessary to accurately model the electrical and thermal spreading resistance, however simulation of 3D structures is somewhat slow. The 3D electrical and thermal resistances can be determined for various geometries and used in an effective one-dimensional (1D) model[8] to obtain faster results when several parameters are to be varied. Thermionic cooling is considered throughout to make fair comparisons to experimental data, however the results are equally applicable to thermoelectric thin film coolers.

a)Electronic mail: labounty@opto.ucsb.edu

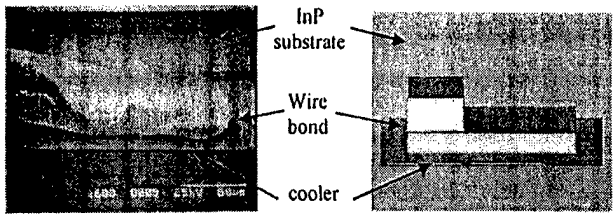


FIG. 1. (a) Scanning electron micrograph of a thin film thermionic cooler (InGaAs/InGaAsP/InGaAs 0.3 μm/1 μm/0.5 μm) and (b) the simulated structure. The uppermost wire bond length is scaled to reduce the element count in the mesh.

II. NONIDEAL EFFECTS

The most evident nonideal effect for thin film coolers is the substrate and package thermal resistance. Qualitatively speaking, if the distance between the cooling and heating regions is several orders of magnitude smaller than the distance between the heating and heat sink regions, most of the heat will flow back to the cold side of the device if the thermal conductivities of the thin film and substrate are comparable. Figure 2(a) shows simulation results of substrate thermal resistance versus thickness for various substrate materials. The simulation is performed assuming a rectangular etched mesa (5000 μm^2) thin film cooler on a semi-infinite plane substrate. The boundary conditions imposed assume the sides and top of the substrate to be adiabatic and the bottom isothermal. A uniform heat load on top of the mesa is assumed. The relatively good fit to a $\ln(x)$ function is indicative of the thermal spreading in two and three-dimensional heat flow.[9,10] Below 15 μm, the heat flow becomes dominantly one dimensional and the thermal resistance scales linearly with substrate thickness. The logarithmic fitting function of the thermal resistance and the corresponding coefficients can be given as

$$R_{\text{thermal}}(h) = a_i \ln(h) + b_i, \tag{1a}$$

$$a_i = \frac{1167}{\beta_i[W/mK]}; \quad b_i = \frac{-1445}{\beta_i[W/mK]}, \tag{1b}$$

where R_{thermal} has units of [K/W], h is the substrate thickness in microns, and a_i and b_i are the fitting coefficients corresponding to the thermal conductivity of the substrate (β_i).

Clearly it is beneficial to use substrates with high thermal conductivity and a minimum thickness. However, thin film coolers that are grown on typical substrate materials (InP, Si, etc.) can be lapped only by a limited amount before the material begins to warp and become severely fragile. This warping occurs with InP substrates when they are lapped below 100 μm, which for the device size simulated corresponds to a thermal resistance of 57 K/W. Using a 1 μm film of InGaAsP for the cooler ($\beta = 3.3$ W/m K) gives a thermal resistance of 61.6 K/W showing that roughly half the heat flows back to the cold side of the device.

The area of the device also affects the thermal resistance. Figure 2(b) shows simulation results of thermal resistance (points on lower curve) for various device areas assuming a 125-μm-thick InP substrate. The upper curve represents a

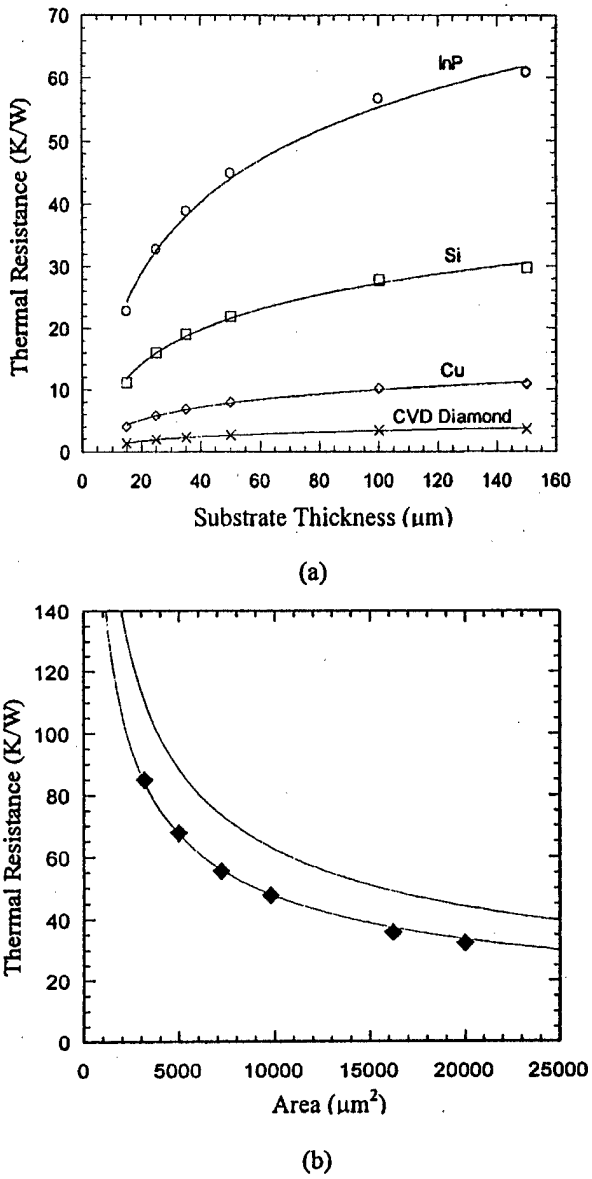


FIG. 2. (a) Thermal resistance vs substrate thickness for a device area of 5000 μm^2 on various substrates: InP($\beta = 71$ W/m K), Si($\beta = 145$ W/m K), Cu($\beta = 398$ W/m K), and chemical vapor deposition diamond($\beta = 1200$ W/m K). The points correspond to 3D simulation results, and the solid curves are the theoretical $a_i \cdot \log(x) + b_i$ curve fits. (b) Thermal resistance vs device area. The points are simulation results assuming a 125-μm-thick InP substrate and the upper solid curve is the theoretical plot assuming the substrate is an entire half space.

theoretical expression that assumes a disk heat source on a half space denoting a purely three-dimensional heat flow

$$R_{\text{thermal}} = \frac{\sqrt{\pi}}{4\beta\sqrt{A}}, \tag{2}$$

where β is the thermal conductivity of the half space and A is the area of the disk.[9] The simulation results are fitted with this same expression (solid line), and the resulting expression is equivalent to Eq. (2) multiplied by a reduction factor of 0.761. As the substrate thickness is increased, the simulation results approach that of the theoretical expression.

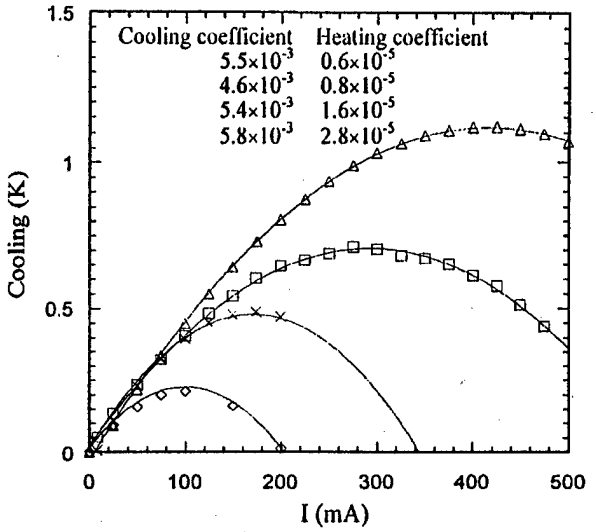


FIG. 3. Measured cooling for a 1-μm-thick cooler (InGaAs/InGaAsP superlattice) with various packages. As the packaging is improved the heating coefficient is reduced resulting in greater absolute cooling. The cooling (linear) and heating (quadratic) coefficients correspond to a second order polynomial fit of ΔT vs I. All temperatures are relative to the value at zero current with a heat sink temperature of 20 °C.

Instead of working with thin substrates that have a low thermal conductivity, switching to a substrate with a very high thermal conductivity would be advantageous. Several methods for transferring thin films to surrogate substrates with better thermal properties are possible, but each has potential problems. Epitaxial lift-off and grafting of epitaxial films has been discussed in the literature[11–14] for various semiconductor material systems, but neither the electrical nor thermal contact resistance of the interface has been characterized. Another possibility is to use flip chip bonding techniques to accomplish the substrate transfer, however thermal expansion mismatch between the thin films and the new substrate can cause problems in material quality.

Packaging of the single element thin film coolers was necessary to attain significant cooling since simply probing the cold side of the device presented too large a heat load. Using wire bonds from the package to the device alleviated some of this problem, but a significant heat load from the wire bond remained as well as the additional thermal resistance of the package between the hot side of the cooler and the heat sink. An optimum wire length can be determined by considering the trade-off between Joule heating in the wire and thermal conduction away from the cold junction. Ultimately when the thin film coolers are integrated with real devices or are packaged in a conventional TE configuration (n- and p-type legs electrically in series, thermally in parallel—see, for example, Ref. 15), the issue of wire bonding will be less of a concern. Figure 3 shows the increase in InGaAsP thin film cooler performance for four generations of packages. The details of the cooler will be discussed later. Generally in each successive package generation the optimum wire bond length is approached and the heat sinking is improved by choosing package materials with higher thermal conductivities. The best cooling shown in Fig. 3 is for a package made from a 350-μm-thick silicon wafer and using a wire bond length of 250 μm. The sample was mounted with a 4-μm-thick tin solder layer. Qualitatively, when the packaging improves the linear coefficient that represents thermionic and thermoelectric cooling effects remains relatively constant while the quadratic coefficient that represents Joule heating is reduced. This results in greater net cooling. One should note that, due to a relatively high contact resistance, the maximum current that can be applied to the thin film device is limited and thus we are still in a linear transport regime. With improvements in nonideal effects and by using asymmetric barriers, one can benefit fully from the thermionic cooling process by going into the nonlinear transport regime and thus increasing the cooling power density.[16]

Another important nonideal effect for thin film coolers is the electrical contact resistance. Because the InGaAsP-based layers have a low electrical resistivity ($\rho = 1.17 \times 10^{-3}\ \Omega$ cm) the contact resistivity usually dominates. This additional Joule heating occurs very close to the cooling region of the device and must be minimized to attain any appreciable cooling. In order to determine the quality of our ohmic contacts, contact resistivity studies were performed with four point probe measurements on transmission line model test patterns.[17] The lowest measured value of specific contact resistance was roughly $5.5 \times 10^{-7}\ \Omega\ \text{cm}^2$ for alloyed Ni/AuGe/Ni/Au contacts to 0.5-μm-thick $n+$InGaAs ($2 \times 10^{19}\ \text{cm}^{-3}$). By studying various annealing conditions, it was determined for the InGaAs alloyed contacts that rapid heating and cooling produced the lowest contact resistance in agreement with previous literature.[18,19] The alloying depth should also be considered when designing thin film coolers. This depth is equal to the amount of semiconductor that intermixes with the contact metal and is roughly 0.1–0.2 μm in most III–V systems.[18–20] In our designed thin film coolers, the top $n+$ or $p+$ contact regions are as short as possible for low resistance but safely greater than the expected alloying depth. Nonalloyed contacts would circumvent this situation, however they typically have a higher contact resistance and the trade-off between the contact resistance and the resistance of a thinner contact region should be considered. The effects of contact resistance on cooler performance are discussed further in the next section.

III. COMPLETE DEVICE SIMULATION

Complete three-dimensional device structures with all nonideal effects included are simulated to fit experimental data and determine which areas of the thin film cooler designs need to be improved. These simulations model self-consistently the thermal and electrical operation. Once the simulation is in agreement with experimentally measured temperature profiles, particular nonideal effects can be removed in succession and the dominating ones determined.

Measured cooling for a 1-μm-thick, $50 \times 100\ \mu\text{m}^2$ thin film cooler (best device from Fig. 3) is shown with the simulated curve in Fig. 4 (curve 1). To check for the presence of thermionic cooling, reference samples with no barrier (InGaAs only) were tested. The reference sample displayed a cooling that was a factor of 2 smaller. Since the reported thermal conductivity of quaternary InGaAsP is only 20%–

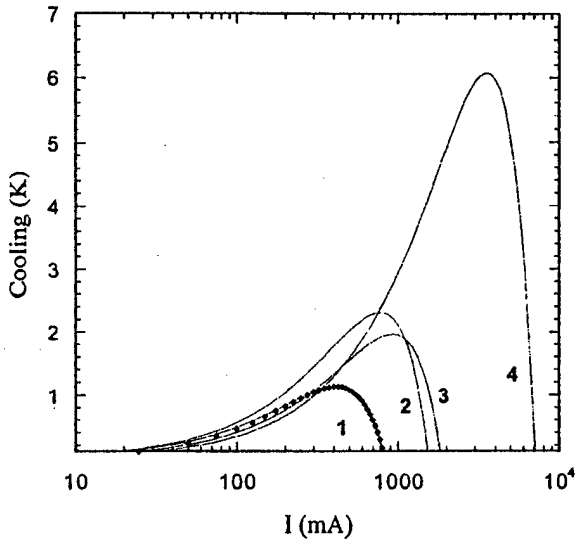


FIG. 4. Experimental and simulation results for the 1-μm-thick cooler shown in Fig. 3. The points and curve 1 correspond to the simulated fit of the experimental data with all nonideal effects. Curves 2 and 3 are the repeated simulation results when the contact resistance and substrate thermal resistance are taken away, respectively. Curve 4 corresponds to both nonideal effects removed.

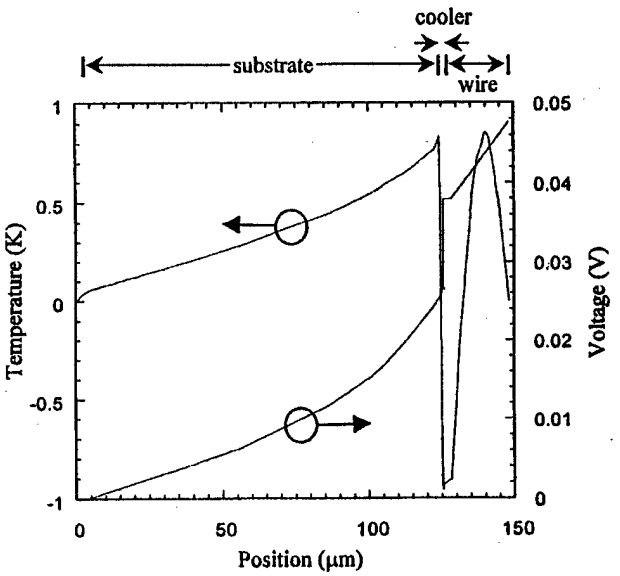


FIG. 5. Simulated temperature and voltage profiles along a path through the substrate, cooler, and wire bond. The actual wire bond length was 400 μm, but was scaled by 20× to reduce the number of elements in the 3D simulation.

30% less than InGaAs, one expects that thermionic cooling plays a significant role in the device operation. Both the thermionic and thermoelectric cooling effects were included in the simulation and the thermal conductivity of the InGaAs/InGaAsP superlattice was taken to be 5 W/m K. With an accurate model of the device in hand, the contact resistance was set to zero and the simulation was repeated (curve 2). The maximum cooling temperature increased from 1.14 to 2.3 °C. This was repeated with the contact resistance reset to its original value and the 120-μm-thick InP substrate replaced with a 10 μm copper substrate resulting in a similar improvement in performance (curve 3). The simulation was then again performed with the top wire bond removed (not shown), resulting only in a very small increase in cooling. Finally, the simulation was repeated once more with both the contact resistance removed and the copper substrate resulting in a maximum cooling temperature of 6.07 °C or a cooling power density of 1821 W/cm^2 (curve 4). Therefore in this particular device structure, both the contact resistance and the substrate thermal resistance will need to be reduced to see substantial improvement in performance.

It is insightful to examine the temperature and voltage profiles along a line through the 3D structure. Figure 5 shows simulation results following a path from the bottom of the substrate, traveling through the cooler, and continuing to the end of the wire bond. The current source is applied between the end of the wire bond and the bottom substrate plane, and an ideal heat sink is assumed at these two points. The wire bond (25 μm diameter) was 400 μm long, but was scaled by a factor of 20 to reduce the number of elements in the 3D simulation. A linear voltage drop and quadratic temperature distribution is observed along the wire bond as expected for one-dimensional behavior. At the cooler, a steep temperature gradient is observed across the 1-μm-thick film indicating a large heat flux. There is also a sharp voltage drop located at the interface of the metal and semiconductor corresponding to the contact resistance which was assumed to be $6 \times 10^{-7}\ \Omega\ \mathrm{cm}^2$. This voltage drop is nearly one quarter the total voltage across the entire structure. Inside the substrate, the voltage and temperature profiles can be seen to drop off nonlinearly due to the three-dimensional spreading. Finally, near the bottom of the substrate there is another change in the slope due to a 4-μm-thick tin solder layer.

From the three-dimensional modeling, it is possible to extract the important thermal and electrical spreading resistances similar to Fig. 2. Using these simulated values it becomes possible to construct an accurate one-dimensional model. Figure 6 shows the 1D model and boundary conditions applied to the heat flux equation

$$\kappa_x A_x \frac{d^2T}{dx^2} = \frac{I^2}{\sigma_x A_x}, \tag{3}$$

which relates the temperature gradient to the Joule heat generation where σ_x, κ_x, and A_x, are the electrical conductivity, thermal conductivity, and area, respectively, defined in each of the three regions. The right side of Fig. 6 represents the equivalent circuit model with the thermionic (Q_{TI}) and thermoelectric (Q_{TE}) cooling and heating sources as well as contact resistance (Q_C). The thermionic cooling and heating occur on their respective sides of the thin film, however only one thermoelectric source is included at the upper metal-semiconductor interface since the bottom semiconductor-metal interface occurs at the heat sink. Equation (2) is valid in each region, and by integrating twice with respect to position an expression for temperature with two unknown constants results, giving a total of six unknowns. Four boundary conditions exist for the temperature since it must be continuous across all three regions. Here thermal boundary resistances are neglected. The last two boundary conditions are

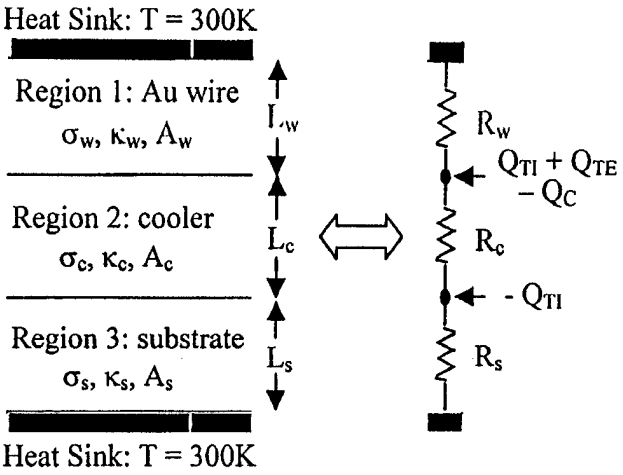


FIG. 6. One-dimensional model and boundary conditions. An electrical conductivity (σ), thermal conductivity (κ), and area (A), are defined in each region. The equivalent circuit model is shown on the right with the arrows indicating sources or sinks of heat flux. Q_{TI} refers to thermionic heating/cooling, Q_{TE} to thermoelectric cooling, and Q_C to heat generation by contact resistance.

obtained from the discontinuity of the heat conduction ($\kappa A dT/dx$) by the heat generation and absorption on either side of the cooler, resulting in six boundary conditions and six unknowns. The solution was then manipulated and plotted in MATLAB.

Figure 7 compares the results of the 3D and 1D models. The two simulations were found to be in good agreement over various changes in parameters. The largest difference can be seen in the temperature profile through the substrate where the 1D and 3D distributions are inherently different. The other minor discrepancy is in the temperature distribution of the wire that arises from the additional thermal spreading at the wire-cooler interface. This effect is more difficult to include in the 1D model.

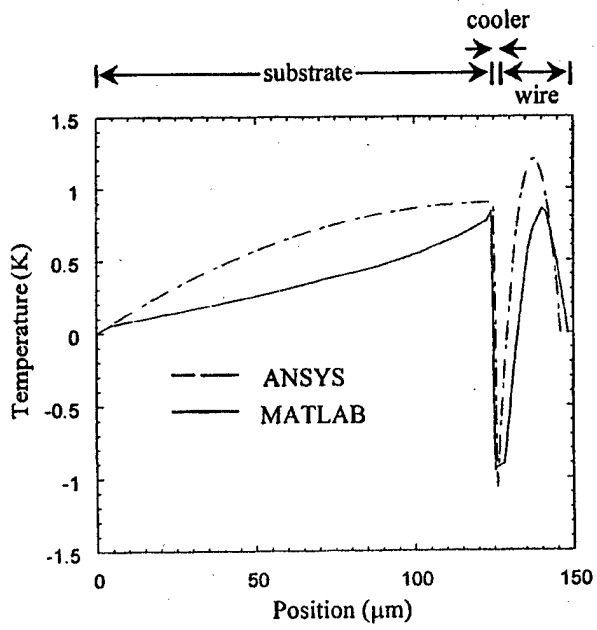


FIG. 7. Comparison of simulated temperature profiles through the substrate, cooler, and wire bond for the three-dimensional (ANSYS) and one-dimensional (MATLAB) models. The zero temperature axis is relative to 300 K.

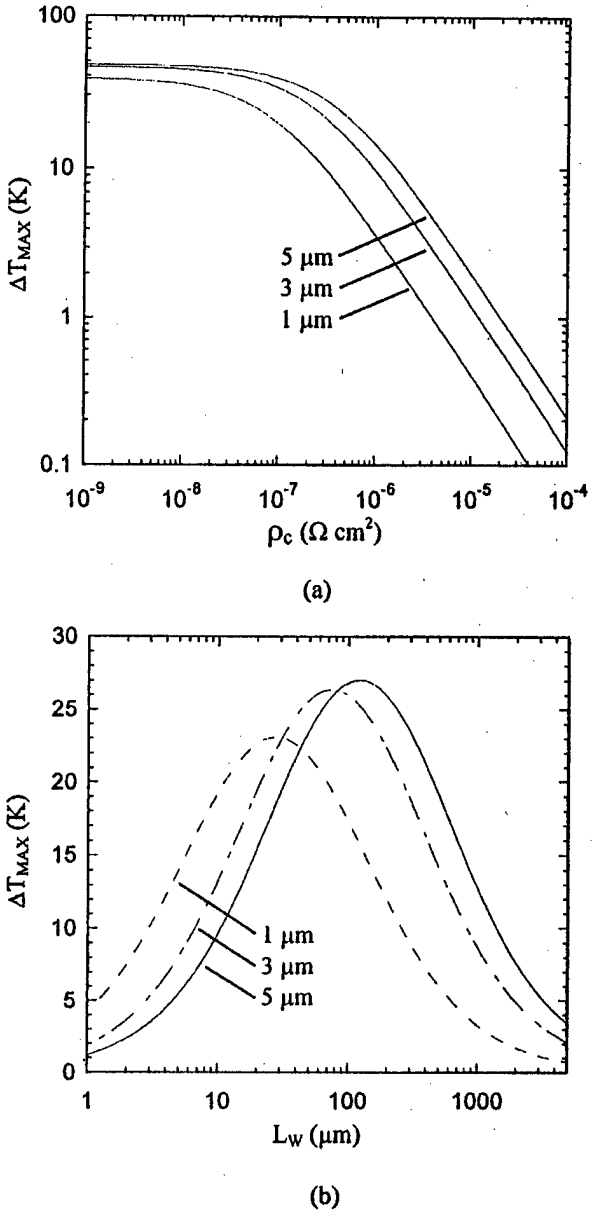


FIG. 8. (a) Simulation of maximum cooling vs contact resistivity for various cooler thicknesses. The Au wire bond was 50 μm long and 50 μm in diameter and zero heat conduction though the wire was assumed as in a p and n cooler configuration. (b) Simulation of single stage maximum cooling vs Au wire bond length (25 μm diameter wire) where now heat conduction through the wire is considered. The contact resistance was assumed to be 10^{-8} Ω cm^2. The substrate thermal resistance was 4 K/W in each case.

The advantage of the one-dimensional model is the speed with which many device parameters can be varied and their affects determined. Figure 8(a) shows a simulation of maximum cooling versus contact resistivity for various cooler thicknesses. To observe just the limitation of contact resistance, the other nonideal effects were minimized to reasonable values. The substrate thermal resistance was assumed to be 4 K/W which from Fig. 2 could represent either a thin 15 μm Cu substrate or an arbitrarily thick diamond

substrate. The wire bond was assumed to be 50 μm long and 50 μm in diameter with no heat conduction through the wire as would be in a p- and n-type conventional thermoelectric configuration. The maximum cooling drops off significantly for a contact resistivity above $10^{-7}\ \Omega\ \mathrm{cm}^2$. Also, thicker devices cool more when the contact resistance is substantial. The reason being that the optimum current for the cooler performance scales as one over the device thickness. A lower current results in less Joule heating from the contact resistance. Optimum currents for the thicker devices were less than 1 A, while the thinner devices required as much as 5 A. While the thicker devices do attain higher absolute cooling in this case, the cooling power remains approximately unchanged.

In applications for which single element coolers are monolithically integrated with electronic or optoelectronic devices, it is necessary to have an external wire bond connected directly to the cold side of the cooler. The question arises as to whether useful cooling can still be achieved in this configuration. Figure 8(b) shows a simulation of maximum cooling versus wire bond length (25 μm diameter). In this case heat conduction through the wire is considered, and the contact resistivity is assumed to be $10^{-8}\ \Omega\ \mathrm{cm}^2$ in order to study the effects of the wire only. For a given cooler thickness, there exists an optimum wire length resulting from a trade-off between Joule heating in long wires and heat conduction from heat sink to the cold junction in short wires. At longer wire lengths the thicker devices cool better by the same argument made for contact resistance. That is thicker coolers are optimized at a lower current and hence less Joule heating occurs. The cooling power of the thicker devices is reduced somewhat, however. For shorter wire lengths the thinner cooler performs best. Since the thinner cooler has a smaller thermal resistance between cold and hot junctions, more heat conduction occurs across the cooler than through the wire.

The temperature dependence of material properties and cooling amount are not taken into account in the above models. For small values of cooling and heating on the order of 5–10 °C, this is a reasonable assumption. However, these effects will need to be considered in future simulations.

IV. CONCLUSIONS

Important parameters and nonideal effects in thin film coolers have been discussed through experimental and simulation results. A three-dimensional finite element simulation has been developed and used to determine the dominating nonideal mechanisms for thin film coolers and the impact of changing device characteristics. A one-dimensional simulation was also developed using three-dimensional spreading resistance values obtained from the three-dimensional model. Contact resistance, finite thermal resistance of substrate and heat sink, and heat generation in wire bonds have all been identified as limitations in thin film cooler performance. Experimental results in thin film thermionic emission coolers have demonstrated cooling by 1.1–1.2 °C at room temperature with cooling power densities of several 100 W/cm^2, and simulations have predicted cooling of 20–30° with cooling power densities of several 1000 W/cm^2 for more optimized structures and packaging.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research under Contract No. 442530-25845, and the Army Research Office through the DARPA/HERETIC program under Contract No. 442530-23002.

[1] L. Rushing, A. Shakouri, P. Abraham, and J. E. Bowers, Proceedings of the 16th International Conference on Thermoelectrics, Dresden, Germany, August 1997, pp. 646–649.
[2] T. A. Corser, 41st Electronic Components and Technology Conference, Atlanta, GA, May 1991, pp. 150–156.
[3] A. Shakouri, C. LaBounty, P. Abraham, J. Piprek, and J. E. Bowers, Mater. Res. Soc. Symp. Proc. **545**, 449 (1999).
[4] A. Shakouri, C. LaBounty, J. Piprek, P. Abraham, and J. E. Bowers, Appl. Phys. Lett. **74**, 88 (1999).
[5] G. Chen, Phys. Rev. B **57**, 14958 (1998).
[6] P. Hyldgaard and G. D. Mahan, Phys. Rev. B **56**, 10754 (1997).
[7] ANSYS Release No. 5.5.3 (1999) by Swanson Analysis Systems, Inc., Houston, PA.
[8] MATLAB Release No. 5.0.0.4073 (1996) by The Math Works, Inc.
[9] L. Coldren and S. Corzine, in *Diode Lasers and Photonic Integrated Circuits* (Wiley, New York, 1995), pp. 55–57.
[10] F. Masana, IEEE Trans. Compon., Packag. Manuf. Technol., Part A **19**, 539 (1996).
[11] J. C. Fan, C. P. Lee, C. M. Tsai, S. Y. Wang, and J. S. Tsang, J. Appl. Phys. **83**, 466 (1998).
[12] C. B. Morrison, R. L. Strijek, and J. H. Bechtel, Proc. SPIE **3005**, 120 (1997).
[13] A. Yi-Yan, W. K. Chan, T. S. Ravi, T. J. Gmitter, R. Bhat, and K. H. Yoo, Electron. Lett. **28**, 341 (1992).
[14] E. Yablonovitch, D. M. Hwang, T. J. Gmitter, L. T. Florez, and J. P. Harbison, Appl. Phys. Lett. **56**, 2419 (1990).
[15] D. M. Rowe, *CRC Handbook of Thermoelectrics* (CRC, New York, 1995).
[16] A. Shakouri and J. E. Bowers, Appl. Phys. Lett. **71**, 1234 (1997).
[17] G. K. Reeves and H. B. Harrison, IEEE Electron Device Lett. **EDL-3**, 111 (1982).
[18] M. Ogawa, J. Appl. Phys. **51**, 406 (1990).
[19] G. Y. Robinson, Solid-State Electron. **18**, 331 (1995).
[20] G. K. Reeves, P. W. Leech, and H. B. Harrison, Solid-State Electron. **38**, 745 (1995).

SiGeC/Si superlattice microcoolers

Xiaofeng Fan,[a] Gehong Zeng, Chris LaBounty, and John E. Bowers
Department of Electrical and Computer Engineering, University of California, Santa Barbara, California 93106

Edward Croke
HRL Laboratories, LLC, Malibu, California 90265

Channing C. Ahn
California Institute of Technology, Pasadena, California 91125

Scott Huxtable and Arun Majumdar
Department of Mechanical Engineering, University of California, Berkeley, California 94720

Ali Shakouri
Baskin School of Engineering, University of California, Santa Cruz, California 95064

(Received 8 November 2000; accepted for publication 24 January 2001)

Monolithically integrated active cooling is an attractive way for thermal management and temperature stabilization of microelectronic and optoelectronic devices. SiGeC can be lattice matched to Si and is a promising material for integrated coolers. SiGeC/Si superlattice structures were grown on Si substrates by molecular beam epitaxy. Thermal conductivity was measured by the 3ω method. SiGeC/Si superlattice microcoolers with dimensions as small as 40×40 μm^2 were fabricated and characterized. Cooling by as much as 2.8 and 6.9 K was measured at 25 °C and 100 °C, respectively, corresponding to maximum spot cooling power densities on the order of 1000 W/cm^2. © *2001 American Institute of Physics.* [DOI: 10.1063/1.1356455]

Thermoelectric (TE) refrigeration in a solid-state active cooling method with high reliability. Bi_2Te_3-based TE coolers are widely used for cooling and temperature stabilization of microelectronic and optoelectronic devices, but their processing is a bulk technology and is incompatible with integrated circuit fabrication process. Solid-state coolers monolithically integrated with microelectronic and optoelectronic devices are an attractive way to achieve compact and efficient cooling. It can lower the cost of fabrication and packaging, and can selectively cool individual key devices instead of the whole chip. However, the thermoelectric figure of merit (ZT) is quite low for most of the semiconductors used in microelectronics and optoelectronics. This makes it difficult to get high cooling performance. Recently heterostructure thermionic and superlattice coolers have been proposed, and theoretical calculations show that large improvements in ZT can be achieved and efficient refrigeration becomes possible with coolers made of conventional semiconductor materials.[1–8]

More recently, a factor of seven enhancement of ZT relative to bulk Si was measured for a Si/Ge superlattice.[9] SiGe/Si superlattice coolers have also been demonstrated with a maximum cooling of 7.2 K at 150 °C.[10,11] Due to the larger lattice constant of germanium compared to silicon (4.2%), Ge and SiGe grown on silicon are compressively strained, thus buffer layers are required for thick Si/Ge and SiGe/Si superlattice layers. This increases the cost of material growth and the complexity of integration with Si-based devices. By adding a small amount of carbon into the SiGe material system, strain can be adjusted due to the small lattice constant of carbon. By properly selecting the Ge and C ratio, SiGeC can be lattice matched to silicon, and thick SiGeC or SiGeC/Si superlattice can be directly grown on Si without strain. Furthermore, while most of the band offset between SiGeC and Si lies in the valence band, its conduction band offset is larger than that between SiGe and Si.[12,13] This makes it possible to use thermionic emission to enhance the TE cooling for both *n*- and *p*-type SiGeC/Si materials.[1–4] In this letter, we report the experimental results on SiGeC/Si superlattice microcoolers. Superlattice structures can improve the cooler performance by reducing the thermal conductivity between the hot and the cold junctions[14,15] and by selective emission of hot carriers above the barrier layers in the thermionic emission process. Si-based microelectronic devices can be monolithically integrated with these coolers to achieve better performance and reliability.

The SiGeC/Si superlattice sample was grown in a Perkin–Elmer Si molecular beam epitaxy (MBE) system capable of codepositing Si, Ge, and C onto 5 in. Si substrates. Solid Si, Ge, and C were evaporated through the use of e-beam sources controlled by electron impact emission sensors (Si, Ge) and by monitoring atomic mass unit 36 (C_3) with a quadrupole mass spectrometer (C). Prior to loading into the MBE system, 125 mm diameter, (001)-oriented Si substrate, doped to <0.005 Ω cm with As, was stripped using 5% hydrogen fluoride. After rinsing with deionized water, the wafer was loaded in the MBE chamber. In order to remove any remaining oxide and to prepare the sample for epitaxial growth, the substrate was heated to 850 °C and exposed to a 0.1 Å/s Si flux for 30 s.

The structure consisted of a 2 μm thick $Si_{0.89}Ge_{0.10}C_{0.01}$/Si superlattice (100 periods, each sublayer

[a]Electronic mail: fan@opto.ucsb.edu

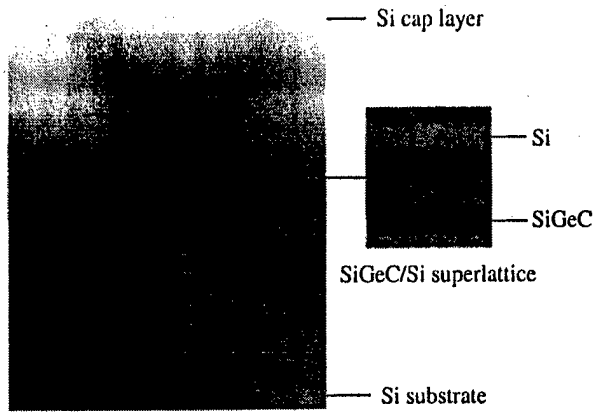


FIG. 1. Cross sectional transmission electron microscopy image of the MBE grown SiGeC/Si superlattice cooler sample. The top is 100 nm Si cap layer; the middle is 2 μm superlattice of 100×(10 nm $Si_{0.89}Ge_{0.10}C_{0.01}$/10 nm Si); the bottom is the Si substrate. An enlarged superlattice image is shown on the right-hand side.

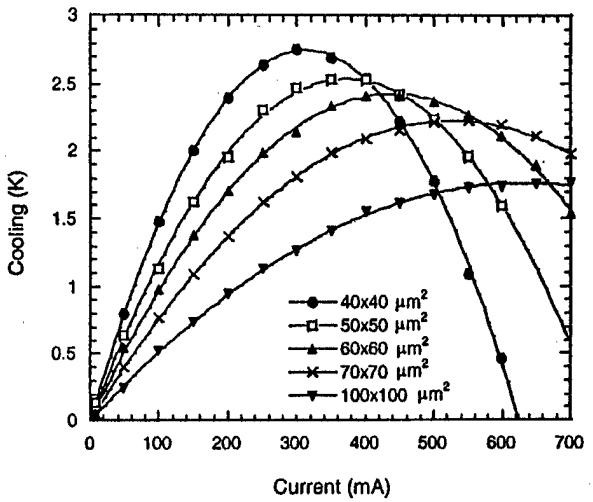


FIG. 3. Meaured cooling for various SiGeC/Si cooler sizes at 25 °C heat sink temperature. The cooler sizes are 40×40 μm^2, 50×50 μm^2, 60×60 μm^2, 70×70 μm^2, and 100×100 μm^2.

10 nm in thickness) grown at 500 °C, lattice matched to the Si substrate. The superlattice was doped with Sb to approximately 2×10^{19} cm^{-3}. Finally, the sample was capped with 100 nm Si:Sb, grown at 460 °C and doped to approximately 1×10^{20} cm^{-3}. A cross sectional transmission electron microscopy image of a grown SiGeC/Si superlattice cooler sample is shown in Fig. 1.

The thermal conductivity is a key parameter for thermoelectric materials. The TE device performance increases with a decrease in thermal conductivity. Thin films and nanostructures have been used to reduce the thermal conductivity via acoustical phonon confinement and interface scattering.[14–17] The 3ω method was used to measure the cross-plane thermal conductivity of the $Si_{0.89}Ge_{0.10}C_{0.01}$/Si superlattice.[18] The result is 0.085 W/cm K, which is over one order lower than that of Si (1.5 W/cm K).

The processing of SiGeC/Si microcoolers is compatible with an integrated circuit fabrication process. The schematic diagram of the cooler device structure is shown in Fig. 2. Microcoolers were thermally isolated by dry etching mesa structures down to the n^+ Si substrate. Metallization was made on the mesa and the Si substrate for cathode and anode contacts, respectively. The main part of the cooler structure is the 2 μm thick superlattice. Its low electrical resistance requires low contact resistance for optimum device performance.[19] Ti/Al metallization was used for ohmic contact. This was followed by annealing at 450 °C for 5 s. Specific contact resistivity of 1.5×10^{-7} Ωcm^2 was measured by transfer length method.

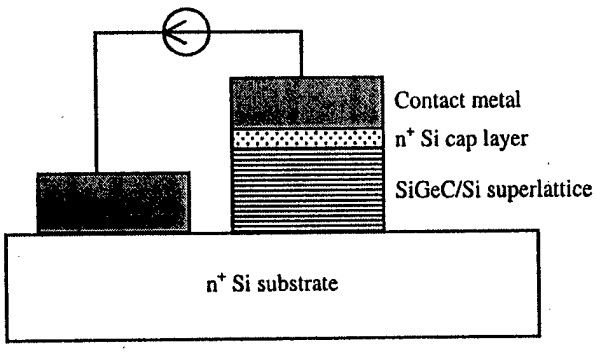


FIG. 2. Schematic diagram of SiGeC/Si microcoolers (not to scale).

SiGeC/Si superlattice microcoolers with various mesa sizes ranging from 40×40 to 100×100 μm^2 were fabricated on one wafer. They were tested on a temperature-controlled heat sink, which was set at a constant temperature during device testing. Device cooling was measured with micro thermocouples on top of the device and was relative to the values at zero current. Figure 3 displays the measured cooling on top of the devices as a function of current with the heat sink at 25 °C. 2.8 K cooling was obtained for the 40×40 μm^2 devices. The test results show that the maximum cooling temperature increases as the device size decreases. This cannot be explained with ideal thermoelectric or thermionic models, and is due to the three-dimensional nature of current and heat spreading in the substrate. The Si substrate of the cooler devices is 500 μm thick and its thermal resistance for the micro devices is about inversely proportional to the square root of the device area. On the other hand, the thermal resistance of the SiGeC/Si superlattice layer is inversely proportional to the device area. This different size dependence makes the effect of the nonideal heat sink smaller for smaller size devices.

For comparison, Si microcoolers were fabricated on n^+ Si substrates with similar device structure and processing. The results of the 40×40 μm^2 Si devices are shown in Fig. 4 along with those of the SiGeC/Si superlattice coolers of the same size. Over three-fold improvement in maximum cooling is observed for SiGeC/Si superlattice coolers over Si ones.

SiGe is a good TE material for high temperature applications.[20] The SiGeC/Si microcoolers also show better performance at higher temperatures. Figure 5 shows the measured cooling for 50×50 μm^2 SiGeC/Si microcoolers at various heat sink temperatures. The maximum cooling increases from 2.5 K at 25 °C to 6.9 K at 100 °C.

The maximum cooling power density is given by

$$Q=\kappa(\Delta T_{\max}-\Delta T)/d,$$

where κ and d are the thermal conductivity and the thickness

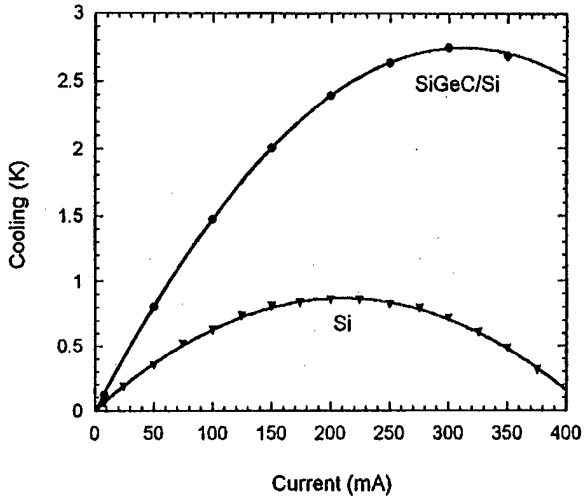


FIG. 4. Measured cooling for 40×40 μm^2 SiGeC/Si and Si coolers at 25 °C heat sink temperature.

of the superlattice, while ΔT_{max} and ΔT are the maximum and actual cooling temperature across the superlattice. At zero ΔT, the coolers have the largest cooling power. Since the majority of the cooling happens over the 2 μm SiGeC/Si superlattice layer, several degrees of cooling corresponds to maximum cooling power densities on the order of 1000 W/cm^2 at zero temperature difference.

The maximum cooling temperature of the SiGeC/Si superlattice microcoolers is limited by the contact resistance, Joule heating and heat conduction from the metal wire connected to the cold junction of the cooler, and the low ZT of the Si substrate. These can be solved by increasing the superlattice thickness, making *n*-type and *p*-type cooler arrays and substrate removal, respectively. Fundamentally, the cooler performance is determined by the TE ZT of the cooler materials. A superlattice gives more freedom in material engineering both electrically and thermally, and enables one to enhance the thermoelectric cooling by thermionic emission,[1-4] quantum confinement,[5,6] carrier pocket engineering,[7,8] and phonon engineering.[16,17] The structure and doping of the SiGeC/Si superlattice studied here have not been optimized yet. With optimized material and device design and packaging, cooling up to tens of degrees is possible. Furthermore, SiGeC/Si superlattices can be lattice matched to Si, and SiGeC/Si microcoolers can be monolithically integrated with Si-based devices to achieve compact and efficient localized cooling.

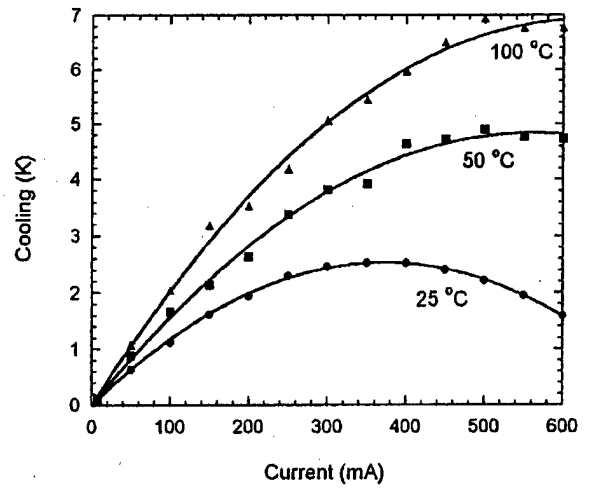


FIG. 5. Meausred cooling for 50×50 μm^2 SiGeC/Si coolers at various heat sink temperatures.

In summary, a lattice matched SiGeC/Si superlattice was grown on Si by MBE and SiGeC/Si superlattice microcoolers were demonstrated. Cooling by as much as 2.8 and 6.9 K was measured at heat sink temperatures of 25 °C and 100 °C, respectively, corresponding to maximum cooling power densities on the order of 1000 W/cm^2.

This work is supported by DARPA HERETIC program and the Army Research Office.

[1] A. Shakouri and J. E. Bowers, Appl. Phys. Lett. **71**, 1234 (1997).
[2] A. Shakouri, C. LaBounty, P. Abraham, J. Piprek, and J. E. Bowers, Mater. Res. Soc. Symp. Proc. **545**, 449 (1999).
[3] G. D. Mahan and L. M. Woods, Phys. Rev. Lett. **80**, 4016 (1998).
[4] C. B. Vining and G. C. Mahan, J. Appl. Phys. **86**, 6852 (1999).
[5] L. D. Hicks and M. S. Dresselhaus, Phys. Rev. B **47**, 12727 (1993).
[6] L. D. Hicks, T. C. Harman, and M. S. Dresselhaus, Appl. Phys. Lett. **63**, 3230 (1993).
[7] T. Koga, X. Sun, S. B. Cronin, and M. S. Dresselhaus, Appl. Phys. Lett. **73**, 2950 (1998).
[8] T. Koga, X. Sun, S. B. Cronin, and M. S. Dresselhaus, Appl. Phys. Lett. **75**, 2438 (1999).
[9] T. Koga, S. B. Cronin, and M. S. Dresselhaus, Appl. Phys. Lett. **77**, 1490 (2000).
[10] X. Fan, G. Zeng, E. Croke, G. Robinson, C. LaBounty, A. Shakouri, and J. E. Bowers, Phys. Low-Dimens. Semicond. Struct. **5**, 1 (2000).
[11] G. Zeng, A. Shakouri, C. LaBounty, G. Robinson, E. Croke, P. Abraham, X. Fan, H. Reese, and J. E. Bowers, Electron. Lett. **35**, 2146 (1999).
[12] H. J. Osten, J. Appl. Phys. **84**, 2716 (1998).
[13] B. L. Stein, E. T. Yu, E. T. Croke, A. T. Hunter, T. Laursen, A. E. Bair, J. W. Mayer, and C. C. Ahn, Appl. Phys. Lett. **70**, 3413 (1997).
[14] S. M. Lee, D. G. Cahill, and R. Venkatasubramanian, Appl. Phys. Lett. **70**, 2957 (1997).
[15] R. Venkatasubramanian, Phys. Rev. B **61**, 3091 (2000).
[16] A. Balandin, A. Khitun, J. L. Liu, K. L. Wang, T. Borca-Tasciuc, and G. Chen, in The 18th International Conference on Thermoelectrics: ICT Symposia Proceedings, Baltimore, Maryland, 29 August–2 September 1999, p. 189.
[17] G. Chen, T. Zeng, T. Borca-Tasciuc, and D. Song, Mater. Sci. Eng., A **292**, 155 (2000).
[18] D. G. Cahill, Rev. Sci. Instrum. **61**, 802 (1990).
[19] C. LaBounty, A. Shakouri, G. Robinson, P. Abraham, and J. E. Bowers, in The 18th International Conference on Thermodynamics: ICT Symposia Proceedings, Baltimore, Maryland, 29 August–2 September 1999, p. 23.
[20] C. B. Vining, J. Appl. Phys. **69**, 331 (1991).

Conclusions: A turbo TCM/FPM scheme has been proposed that can achieve considerable coding gains over turbo TCM 8PSK and conventional TCM/FPM schemes without sacrificing power and bandwidth efficiency. The performance of a DS/SSMA system using turbo TCM/FPM improves significantly over a multipath Rician fading channel. This scheme is attractive for power-limited and band-limited environments.

6 November 2000
Electronics Letters Online No: 20010053
DOI: 10.1049/el:20010053

Xiaoming Peng and E. Gunawan (*School of Electrical and Electronic Engineering, Blk. S1, Nanyang Technological University, Singapore 639798, Republic of Singapore*)

E-mail: egunawan@ntu.edu.sg

References

1 ROBERTSON, P., and WORZ, T.: 'Bandwidth-efficient turbo trellis-coded modulation using punctured component codes', *IEEE J. Sel. Areas Commun.*, 1998, **16**, (2)
2 PADOVANI, R., and WOLF, J.K.: 'Coded phase/frequency modulation', *IEEE Trans. Commun.*, 1986, **COM-34**, pp. 446–453
3 DU, JUN, VUCETIC, B., and ZHANG, LIN: 'Construction of new MPSK trellis coded modulation for fading channels', *IEEE Trans. Commun.*, 1995, **43**, (2), pp. 776–784

High cooling power density SiGe/Si micro-coolers

X. Fan, G. Zeng, E. Croke, C. LaBounty, C.C. Ahn, D. Vashaee, A. Shakouri and J.E. Bowers

SiGe/Si superlattice micro-coolers are investigated experimentally. They can be monolithically integrated with Si-based microelectronic devices to achieve localised cooling and temperature control. Cooling by as much as 4.2K at 25°C and 12K at 200°C was measured on 3μm thick, 60 × 60μm² devices. This corresponds to maximum cooling power densities approaching kW/cm².

Thermoelectric (TE) refrigeration is an active solid-state cooling method with high reliability. TE coolers have been widely used in cooling and temperature stabilisation of microelectronic and opto-electronic devices to achieve precise operational characteristics. Conventional TE coolers are discrete devices with dimensions of the order of millimetres to centimetres, and they generally cool an entire packaged chip. Micro-coolers monolithically integrated with microelectronic and optoelectronic devices provide a more efficient way of temperature regulation by selectively cooling critical components. Compared with traditional TE coolers, integrated coolers are more compact, consume less power, and have a faster response. They can also be used to construct multiple independently temperature-controlled regions on a single chip.

For monolithic integration, conventional semiconductor materials with good TE properties are required. SiGe alloy is one of the best TE materials for high temperature applications. At room temperature, based on the TE properties of bulk SiGe [1], it can give a maximum cooling of over 10K. Although this cooling is less than that of Bi_2Te_3 coolers, SiGe coolers can be monolithically integrated with Si-based microelectronic devices and have many potential applications. Recently, superlattice structures have been proposed to increase the thermoelectric figure of merit ZT beyond that of alloy materials by quantum confinement [2], thermionic emission [3] and carrier pocket engineering [4]. SiGe/Si micro-coolers were first demonstrated on *n*-type samples [5]. In this Letter, we report our experimental work on *p*-type SiGe/Si superlattice micro-coolers with improved performance. This paves the road to fabricate an array of *n*- and *p*-type coolers electrically in series and thermally in parallel and thus achieve larger amounts of cooling.

The micro-cooler structure uses cross-plane electrical transport as shown in Fig. 1. The main part of the cooler is a 3μm strain compensated SiGe/Si superlattice. It consists of 200 periods of (12nm $Si_{0.75}Ge_{0.25}$/3nm Si), doped with boron to about $6 \times 10^{19}cm^{-3}$. The $Si_{0.75}Ge_{0.25}$/Si superlattice has a valence band offset of about 0.16eV [6], and hot holes going over this barrier can produce thermionic cooling [3]. This superlattice was grown using molecular beam epitaxy (MBE). Its average lattice constant is that of $Si_{0.8}Ge_{0.2}$, and a buffer layer is required for it to be grown on a Si substrate. To reduce the material growth time in the MBE system, the buffer layer was grown on a p^+ (001) Si substrate by chemical vapour deposition (CVD) in the form of a graded SiGe structure. The boron doping in the buffer layer is $5 \times 10^{19}cm^{-3}$. Following the superlattice growth, a 0.3μm $Si_{0.8}Ge_{0.2}$ cap layer was grown with a boron doping of $2 \times 10^{20}cm^{-3}$ to allow for a good ohmic contact to the device. Fig. 2 shows a cross-section transmission electron microscopy (TEM) image of the MBE grown SiGe/Si superlattice.

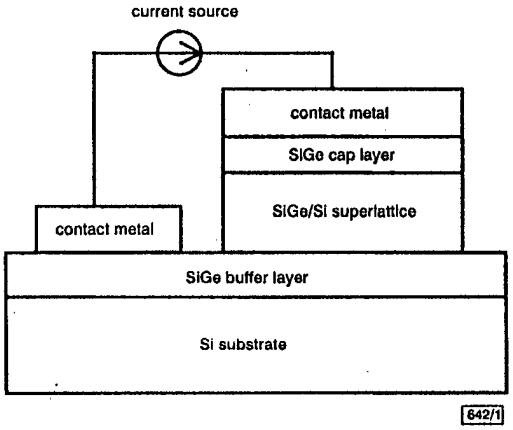


Fig. 1 *Structure of SiGe/Si micro-cooler*

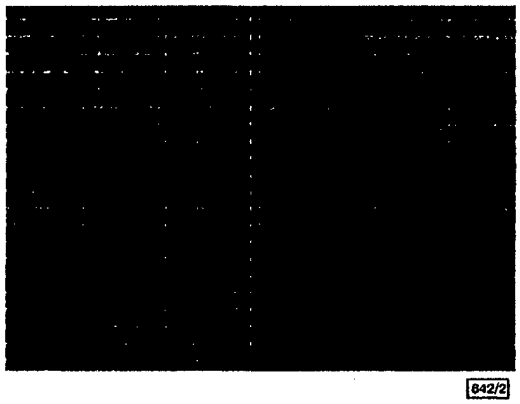


Fig. 2 *TEM image of SiGe/Si superlattice*
Dark lines are 12nm $Si_{0.75}Ge_{0.25}$ layers, light lines are 3nm Si layers

The SiGe/Si micro-coolers are fabricated with standard silicon integrated circuit technology. The cooler device areas were defined by etching mesas down to the SiGe buffer layer. Ti/Al metallisation was made on top of the mesa and on the SiGe buffer layer next to the mesa for top and bottom contacts respectively. SiGe/Si superlattice coolers with mesa sizes ranging from 30 × 30μm² to 150 × 150μm² were fabricated on the same wafer. For comparison, bulk Si micro-coolers were also fabricated on p^+ Si substrates with similar device structures and processing.

The coolers were tested on a heat sink stage. Micro-thermocouples were used to measure the temperature on top of the mesa, and the cooling is relative to the value at zero current. The results for 60 × 60μm² devices are shown in Fig. 3 for both SiGe/Si superlattice and Si coolers at a heat sink temperature of 25°C. Cooling by as much as 4.2K was measured on SiGe/Si superlattice coolers. This is over four-fold improvement compared to the bulk Si coolers.

The SiGe/Si micro-coolers perform better at higher temperatures. Fig. 4 shows the measured cooling on 60 × 60μm² SiGe/Si coolers at various heat sink temperatures. The maximum cooling increased from 4.2K at 25°C to over 12K at 200°C. Since most of

the cooling happens over the 3μm superlattice layer, the corresponding maximum cooling power density is of the order of kW/cm^2. This is much larger than that of conventional bulk TE coolers because of the thin film cooler structure of our devices. The improved cooling at higher temperatures agrees with the thermoelectric properties of SiGe alloy, the TE figure of merit of which, ZT, increases with temperature up to 900°C [1]. The high temperature results of the SiGe/Si micro-coolers suggest potential applications beyond the upper temperature limits of conventional Bi_2Te_3 TE coolers.

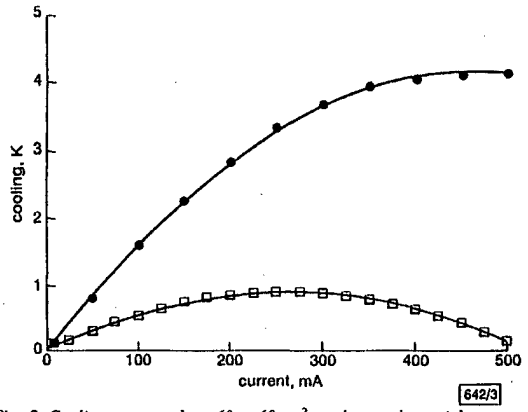


Fig. 3 *Cooling measured on 60 × 60μm² coolers at heat sink temperature of 25°C*

—●— SiGe/Si superlattice
—□— Si

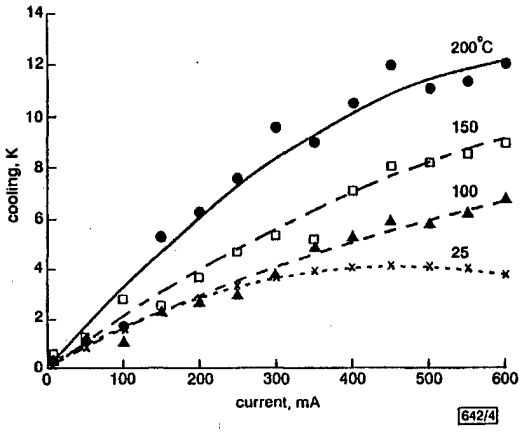


Fig. 4 *Cooling measured on 60 × 60μm² SiGe/Si superlattice coolers at various heat sink temperatures*

Device simulations show that the performance of the micro-coolers can be improved by reducing the contact resistance, reducing the Joule heating and heat conduction from the metal wire connected to the cold junction of the cooler, thinning or removing the Si substrate, and improving the material properties of the superlattice. With optimised device and material deign, cooling up to tens of degree at room temperature is possible. Furthermore, SiGe/Si micro-coolers can be monolithically integrated with Si and SiGe microelectronic devices to achieve compact and reliable localised cooling and active temperature control.

Acknowledgment: This work was supported by the DARPA HERETIC program and the Army Research Office.

7 December 2000
Electronics Letters Online No: 20010096
DOI: 10.1049/el:20010096

X. Fan, G. Zeng, C. LaBounty and J.E. Bowers (*Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, USA*)

E-mail: fan@opto.ucsb.edu

E. Croke (*HRL Laboratories, LLC, Malibu, CA 90265, USA*)

C.C. Ahn (*California Institute of Technology, Pasadena, CA 91125, USA*)

D. Vashaee and A. Shakouri (*Baskin School of Engineering, University of California, Santa Cruz, CA 95064, USA*)

References

1 ROWE, D.M.: 'CRC handbook of thermoeletronics' (CRC Press, New York, 1995)

2 HICKS, L.K., and DRESSELHAUS, M.S.: 'Effect of quantum-well structures on the thermoelectric figure of merit', *Phys. Rev. B*, 1993, **47**, (16), pp. 12727–12731

3 SHAKOURI, A., and BOWERS, J.E.: 'Heterostructure integrated thermionic coolers', *Appl. Phys. Lett.*, 1997, **71**, (9), pp. 1234–1236

4 KOGA, T., SUN, X., CRONIN, S.B., and DRESSELHAUS, M.S.: 'Carrier pocket engineering applied to "strained" Si/Ge superlattices to design useful thermoelectric materials', *Appl. Phys. Lett.*, 1999, **75**, (16), pp. 2438–2440

5 ZENG, G., SHAKOURI, A., LABOUNTY, C., ROBINSON, G., CROKE, E., ABRAHAM, P., FAN, X., REESE, H., and BOWERS, J.E.: 'SiGe micro-cooler', *Electron. Lett.*, 1999, **35**, (24), pp. 2146–2147

6 PEOPLE, R., and BEAN, J.C.: 'Band alignments of coherently strained Ge_xSi_{1-x}/Si heterostructures on ⟨001⟩ Ge_ySi_{1-y} substrates', *Appl. Phys. Lett.*, 1986, **48**, (8), pp. 538–540

Low level and reflection phase noise measurements on a FET

O. Llopis, J.B. Juraver, G. Cibiel and J. Graffeuil

The residual phase noise of a PHEMT device is studied in two unusual configurations: in transmission mode with a low input microwave power on the device, and in reflection mode. Measurements clearly reveal some fundamental aspects of the phase noise generation in this device: the phase noise is a modulation mechanism which still exists in the linear regime and in which the gate reactance fluctuations play an important role.

Introduction: The phase noise generation in an active device has always been a complex and not well understood process. Over past years there have been different, competing, modelling approaches to determine which transistor noise model [1 – 4] or which noise conversion process [5] should be considered. In this Letter, residual phase noise (or open loop phase noise) investigations are shown to be an efficient approach to identify the main noise conversion processes in a microwave field effect transistor (FET). The emphasis is put on low level and reflection mode measurements. The first experiment demonstrates the modulation process involved in the phase noise generation, while the other is the proof of the importance of the gate reactance fluctuations in this mechanism.

Residual phase noise measurement for phase noise modelling: As shown by Leeson in 1966 [6], the active device phase fluctuations $\Delta\phi$ in an oscillator loop are directly converted into frequency fluctuations Δf. Therefore, investigations on Δf can be substituted for investigations on $\Delta\phi$, which feature two strong advantages. The first one deals with the simulation domain: the analysis of a driven circuit (an amplifier) is easier and quicker than the analysis of an autonomous circuit (an oscillator). The second one deals with the experimental domain: the amplifier is a simpler device than an oscillator and its characterisation can be performed in a very precise way, since only two experimental parameters have to be considered, i.e. the bias conditions and the microwave input power. Moreover, phase noise investigations are possible down to the linear regime, in contrast to the oscillator case which is intrinsically nonlinear. The only difficulty is due to the low phase noise levels that must be measured: a low noise phase detector and an appropriate cancellation of the source AM and FM noise contributions are required. Our measurement bench makes use of a two mixer technique, which allows the cancellation of the mixer noise through a cross-correlation process. The input source is a battery-biased 10GHz dielectric resonator oscillator (DRO). The noise floor of the experiment is about –175dBc/Hz at 10kHz offset from

Monolithic integration of thin-film coolers with optoelectronic devices

Christopher LaBounty, MEMBER SPIE
Department of Electrical and Computer Engineering
University of California–Santa Barbara
Santa Barbara, California 93106-9560
E-mail: labounty@opto.ucsb.edu

Ali Shakouri
Jack Baskin School of Engineering
University of California–Santa Cruz
Santa Cruz, California 95064-1077

Patrick Abraham
John E. Bowers
Department of Electrical and Computer Engineering
University of California–Santa Barbara
Santa Barbara, California 93106-9560

Abstract. Active refrigeration of optoelectronic components through the use of monolithically grown thin-film solid-state coolers based on III-V materials is proposed and investigated. Enhanced cooling power compared to the thermoelectric effect of the bulk material is achieved through thermionic emission of hot electrons over a heterostructure barrier layer. These heterostructures can be monolithically integrated with other devices made from similar materials. Experimental analysis of an InP *pin* diode monolithically integrated with a heterostructure thermionic cooler is performed. Cooling performance is investigated for various device sizes and ambient temperatures. Several important nonideal effects are determined, such as contact resistance, heat generation and conduction in the wire bonds, and the finite thermal resistance of the substrate. These nonideal effects are studied both experimentally and analytically, and the limitations induced on performance are considered. Heterostructure integrated thermionic cooling is demonstrated to provide cooling power densities of several hundred W/cm^2. These microrefrigerators can provide control over threshold current, power output, wavelength, and maximum operating temperature in diode lasers. [S0091-3286(00)02511-3]

Subject terms: refrigeration; integrated; thermionic; cooling.

Paper 200084 received Mar. 3, 2000; revised manuscript received June 30, 2000; accepted for publication June 30, 2000.

1 Introduction

Optoelectronic devices such as laser sources, switching/routing elements, and detectors require careful control over operating temperature. This is especially true in current high-speed and wavelength division multiplexed (WDM) optical communication networks. Long-haul optical transmission systems operating around 1.55 μm typically use erbium-doped fiber amplifiers (EDFAs) and are restricted in the wavelengths they can use due to the finite bandwidth of these amplifiers. As more channels are packed into this wavelength window, the spacing between adjacent channels becomes smaller and wavelength drift becomes very important. Temperature variations are the primary cause of wavelength drift, and also affect the threshold current and output power in laser sources. Distributed feedback (DFB) lasers and vertical-cavity surface-emitting lasers (VCSELs) can generate large heat power densities, on the order of kW/cm^2, over areas as small as 100 μm^2.[1] Typical temperature-dependent wavelength shifts for these laser sources are on the order of 0.1 nm/°C. Therefore a temperature change of only a few degrees in a WDM system with a channel spacing of 0.2 to 0.4 nm would be enough to switch data from one channel to the adjacent one, and even less of a temperature change could dramatically increase the crosstalk between two channels. In many optoelectronic applications this temperature dependence is used to actively control the characteristics of the device, as in optical filters[2] or switches.[3,4] In other instances large absolute cooling is desired, as in IR photodetectors.[5] Temperature stabilization or refrigeration is commonly performed with conventional thermoelectric (TE) coolers; however, since integration with optoelectronic devices is difficult,[6] component cost is greatly increased because of packaging. The reliability and lifetime of packaged modules is also usually limited by the TE cooler.[7]

An alternative solution to thermal management needs is to incorporate heterostructure integrated thermionic (HIT) refrigerators with optoelectronic devices.[8,9] These thin-film coolers use the selective thermionic emission of hot electrons over a heterostructure barrier layer. This evaporative cooling occurs because the hot electrons that are on one side of the Fermi energy are emitted. So as to maintain the quasiequilibrium Fermi distribution, lower-energy electrons absorb thermal energy from the lattice at the junction. The emitted electrons then redeposit their energy after passing over the barrier. Since these thin-film coolers can be made with conventional III-V semiconductor materials, monolithic integration with optoelectronics is possible. The result of this integration is an extended component lifetime and a very fast cooling response due to the small thermal mass of the cooler. Furthermore, standard integrated-circuit batch fabrication techniques can be used to manufacture these coolers, whereas TE coolers use a bulk fabrication process. Another advantage of a HIT cooler is the dramatic gain in cooling power density.

To demonstrate monolithic integration of a device with a HIT cooler, a simple InP *pin* diode was grown on top of a 1-μm-thick InGaAs/InGaAsP HIT cooler. Cooling performance was investigated for various currents through the diode, which corresponds to effectively changing the ther-

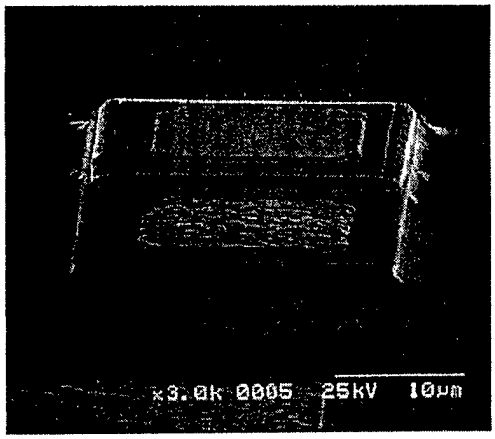


Fig. 1 Scanning electron micrograph (SEM) of a processed device showing the *pin* diode on top of the 1-μm-thick HIT cooler.

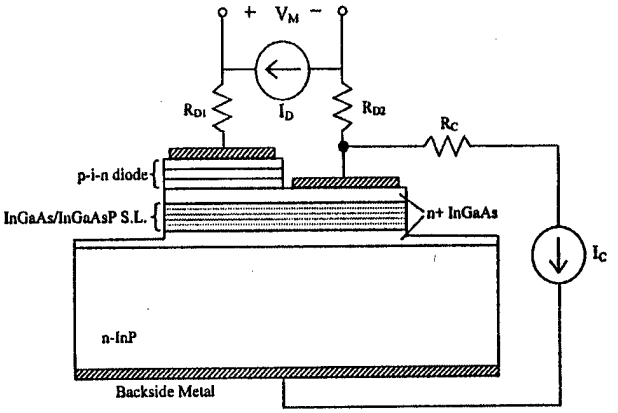


Fig. 2 Setup for measuring the temperature of the integrated diode versus the current through the cooler (I_C) and the diode (I_D). Here R_{D1}, R_{D2}, and R_C are the wire bond resistances.

mal load. After discussing the measurements, the nonideal parasitic effects are identified and examined. These results are then used to predict the effectiveness of HIT-cooler integration with diode lasers.

2 Device Structure

The tested HIT cooler structure consisted of a 1-μm-thick superlattice barrier (25 periods of 10-nm InGaAs and 30-nm InGaAsP, $\lambda = 1.3\,\mu$m) surrounded by n^+ InGaAs cathode and anode layers grown by metal-organic chemical vapor deposition (MOCVD). The cathode and anode layers were 0.3 and 0.5 μm thick, respectively. On top of these layers, a 0.85-μm-thick *pin* diode was grown during the same MOCVD growth. In two wet etching steps, two stacked mesas are defined corresponding to the diode on top of the cooler, as shown in Fig. 1. The cooler mesas ranged in size from 20×40 to $100\times200\,\mu\text{m}^2$, and the diode size was one-half the cooler size. Ti/Pt/Au was used to make ohmic contacts to both the *p*- and *n*-type material. The substrate was thinned to approximately 125 μm before the back-side metal was deposited. The integrated devices were then cleaved, packaged, and wire-bonded for testing.

3 Measurement

The diode serves two purposes in the measurement. By changing the current through the diode we can effectively change the heat load of the cooler. Also, by monitoring the voltage across the diode, we can measure the temperature[10] on the cold side of the cooler, T_C. The temperature sensitivity of the diode near room temperature was determined to be 1.936 mV/°C at a bias of 1 mA. The measurement setup is illustrated in Fig. 2. A constant current (I_D) was sent through the diode, and the diode voltage (V_D) was monitored as the cooler current (I_C) was varied. The resistors of Fig. 2 represent the parasitic wire bonds, which add to the heat load on the device. The measured voltage can be expressed as

$$V_M=(R_{D1}+R_{D2})I_D+V_D+\frac{\rho_c}{A_c}(I_D+I_C), \tag{1}$$

where R_{D1} and R_{D2} are the wire bond resistance, ρ_c is the contact resistivity, and A_c is the area of the metal contact on the cooler. If Eq. (1) is rearranged to solve for V_D as a function of I_C, then the resulting expression is equal to the measured voltage minus some constant values and minus a value that changes with respect to I_C, that is, $\rho_c I_C/A_C$. Therefore, to correctly measure the temperature on top of the device, the contact resistivity must be determined. Once V_D is known, the temperature can be calculated using the temperature dependence of the diode voltage at a constant current as given above.

The thermionic heating and cooling are proportional to the current and can be expressed (in the limit of Boltzmann statistics) as

$$Q_{TI}=\left(\phi_B+\frac{2k_BT}{e}\right)\cdot I, \tag{2}$$

where ϕ_B is the heterojunction barrier height, k_B is Boltzmann's constant, and e is the charge of an electron.[9] Since cooling is linearly proportional to current and Joule heating is proportional to the square of current, the temperature versus current is described by a second-order polynomial. This can best be seen from the expression for overall cooling power:

$$Q_1=Q_{TI}-\frac{\rho d}{A_c}I_C^2-\frac{\beta}{d}\Delta T, \tag{3}$$

where ρ is the electrical resistivity of the cooler, d is its thickness, β is the thermal conductivity between hot and cold junctions, and ΔT is the temperature difference across the cooler. When this expression is solved for ΔT, the terms in the second-order polynomial are apparent.

4 Results

A direct way of measuring the temperature of the device is to use a microthermocouple. Device A from Fig. 3 shows the temperature versus current for the cold side of a $70\times140\,\mu\text{m}^2$ cooler that had the integrated diode removed by selective wet etching. The temperature data was curve-

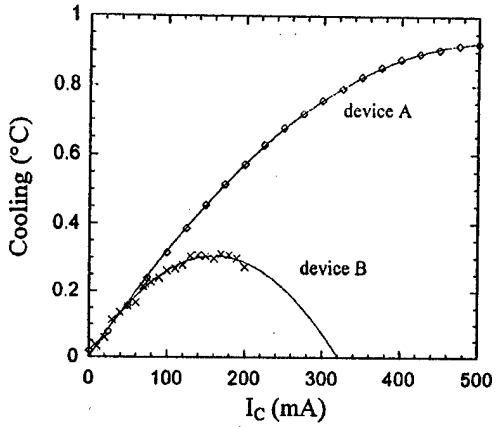


Fig. 3 Measured cooling versus cooler current using a microthermocouple for a well-packaged device (curve 1), and using the integrated *pin* diode with poorer packaging (curve 2). Both device areas were $70\times140\ \mu m^2$.

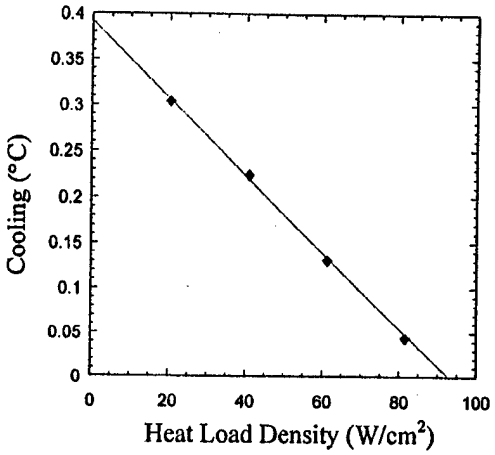


Fig. 4 Cooling versus excess heat load density. The points are experimental data, while the linear curve fit corresponds to Eq. (3).

fitted with a second-order polynomial, resulting in a linear coefficient of 3.53°C/A and a quadratic coefficient of 3.37°C/A^2. Since this cooler structure was identical to the one with the integrated diode (device B), the two devices should have the same linear coefficient. The quadratic coefficient, on the other hand, may differ, since Joule heating is dependent on processing variations such as contact resistance, packaging, and wire bonding. Consequently the contact resistance (ρ_C) can be determined from Eq. (1) by adjusting its value until the linear coefficients match. Using this method, device B from Fig. 3 also shows the temperature versus current for the cold side of a cooler with an integrated diode biased at 1 mA. The performance is much worse, due to poorer packaging and a higher contact resistance. For device B the contact resistance was determined to be $4.4\times10^{-6}\ \Omega\ cm^2$. The contact resistance for device A was measured by the transmission-line model[11] and was estimated to be roughly $5\times10^{-7}\ \Omega\ cm^2$. Knowing the contact resistance, the current through the diode was increased, and the temperature versus cooler current was again measured. Each time the diode current was changed, the temperature sensitivity was recalibrated. By changing the diode current, the heat load is changed. Figure 4 shows the maximum cooling on top of the device as a function of heat load. The dependence of cooling on the heat load density can be described in a similar manner to that of a thermoelectric device,[12]

$$T_C = T_{C\,\max}\cdot\left(1-\frac{Q_l}{Q_{l\,\max}}\right), \tag{3}$$

where T_C is the cold-side temperature, Q_l is the heat load density, and $T_{C\,\max}$ and $Q_{l\,\max}$ are the corresponding maximum values. From Fig. 4 we find that the maximum cooling was 0.39°C, and the maximum heat load density was 93 W/cm^2, considering only heat generation by the diode. In addition to the diode there is a constant heat load due to the two wire bonds attached to the diode (R_{D1} and R_{D2}), and another heat load due to the wire bond to the cooler (R_C), which changes with changing cooler current. In fact, much of the improvement from device B to device A in Fig. 3 can be attributed to the shorter wire bonds. The wire-bond length should not be too short, however, or else the heat conduction through the wire from the cold side to the heat sink will begin to become an issue. An optimum wire length can be determined from the wire's diameter, its electrical and thermal conductivities, and the temperature difference across the wire bond. Figure 5 shows this analysis for a gold wire (25-μm diameter) with various values of current and cooling temperature, assuming that the package side of the wire bond is well heat-sunk. For long wire bonds the Joule heating dominates the cooling power loss, and curves with a similar current approach the same value. For short wire bonds the heat conduction dominates, and curves with a similar temperature difference approach the same value. In between these two extremes there is an optimum wire bond length for a given current and temperature difference.

It is useful to examine the magnitudes of the various sources of heat that are contributing to the total applied thermal load. Figure 6 shows the heat load density versus

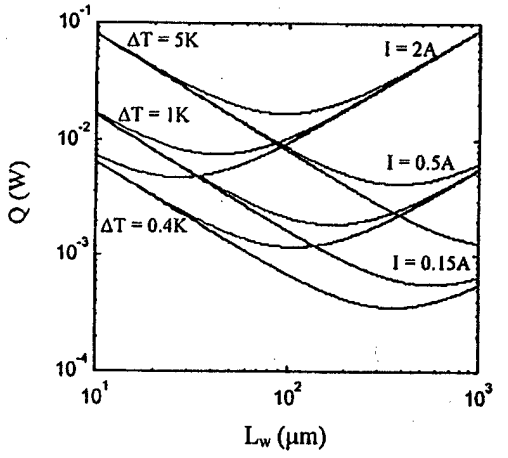


Fig. 5 Cooling power lost versus gold-wire-bond length for various bias currents and temperature differences. The electrical and thermal conductivities of gold were assumed to be $45.5\times10^4\ \Omega^{-1}\ cm^{-1}$ and 3.17 W/cm K, respectively.

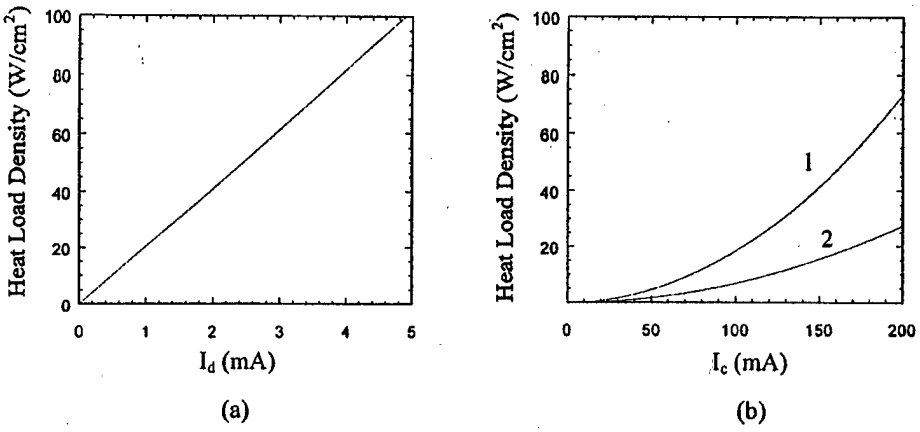


Fig. 6 (a) Heat load density for the diode versus diode current. (b) Heat load density from the wire bond (curve 1) and from the contact resistance (curve 2) versus cooler current.

the respective current bias for the dominating sources of heat. For small temperature differences, it was assumed that approximately one-half the heat generation in the wire bonds arrives at the device, while the other half goes to the heat sink. The heat generation due to the diode contact resistance and wire bonds connected to the diode is not shown, since it was several orders of magnitude smaller due to the small value of the diode current. At the optimum cooler current bias of 160 mA (device B from Fig. 3), there is 47 W/cm^2 of heat load density due to the wire bond connected to the device, and an additional 18 W/cm^2 due to the contact resistance. Using these values, the actual maximum heat load density was 158 W/cm^2.

The smallest coolers yielded the largest absolute cooling. Since a smaller device requires less current to achieve the same temperature gradient, there is less current for parasitic Joule heating from the wire bonds and contact resistance, and the overall cooling is larger. Increasing the heat-sink temperature resulted in a further increase in cooling. Figure 7 shows the maximum cooling increasing from 1.15 to 2.25°C when the heat-sink temperature was raised from 20 to 90°C. One reason for the improvement is the larger thermal spread of carriers at higher heat-sink temperatures, which allows for more carriers to pass over the heterojunction barrier. The other reason is the reduced thermal conductivity of the barrier material, cutting the amount of heat that returns to the cold junction.

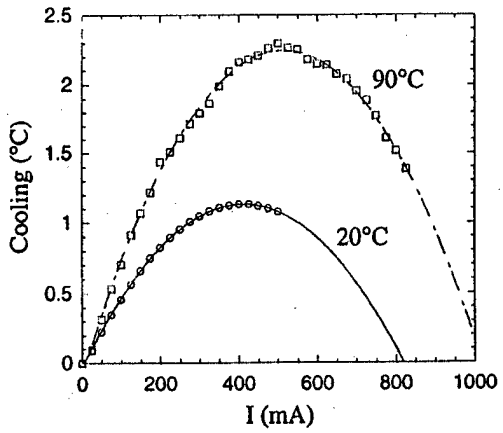


Fig. 7 Measured cooling at heat-sink temperatures of 20 and 90°C.

5 Application to Optoelectronics

The benefits of integrated thin-film cooling for optoelectronic devices can be discussed in three regimes, described as temperature tuning, heat pumping, and absolute cooling. In the first regime the temperature is to be tuned by a moderate amount to control the operating characteristics such as the emission wavelength in a laser diode. In a laser diode operating near or above lasing threshold, most of the heat is generated in the vicinity of the active region due to nonradiative recombination and absorbed radiative recombination.[13,14] Since there is a finite thermal resistance between the active region and heat sink, the temperature of the active region will be greater than that of the heat sink. This thermal resistance is mostly due to the substrate underneath the laser for in-plane geometries, and a combination of the substrate and bottom mirror for vertical cavity geometries. By using the integrated cooler, the active-region temperature may be controlled more precisely and rapidly, as the distance to the cooler can be as small as a few microns and the mass of the cooler is comparable to that of the laser. This is also true in other optoelectronic devices such as filters and switches. Many of these devices make use of Bragg gratings or nonlinear waveguides, where it is important to maintain the temperature at a certain value to achieve the desired phase-matching condition. In many of these cases, the cooler would often be operated in reverse bias to heat as well as cool the device.

The active region in a typical diode laser can reach temperatures more than 70°C above the heat-sink temperature. In this regime the microrefrigerator does not need to cool the laser below the heat-sink temperature, but only needs to provide high heat pumping densities. If the conventional cold side of the cooler is actually hotter than that of the heat sink, then the heat conduction [third term in Eq. (3)] is in the opposite direction and larger cooling power densities are possible. For example, if the active region of a DFB laser with a temperature-dependent wavelength shift of 0.1 nm/°C can be cooled to the heat-sink temperature, then a

tuning of 7 nm should be possible. This wavelength shift could also be used to monitor the temperature of the active region, providing yet another way to characterize the thin-film cooler's performance. Similarly, the output power for a typical DFB laser changes by approximately 0.4 dB/°C. Temperature sensitivity in other kinds of laser diodes is typically much greater. Broad-area Fabry-Perot devices operating in the red and near-IR spectral regions typically have wavelength sensitivities of 0.5 nm/°C,[15] resulting in a larger tunability. Furthermore, if the temperature of the active region could be kept close to that of the heat sink, then the maximum operating temperature of the laser would be equal to that of the heat sink. Keeping the temperature of the active region close to that of the heat sink would also inevitably extend the lifetime of the laser. Finally, with a cooled active region, the carrier leakage out of the separate confinement heterostructure layers would be reduced, resulting in a lower threshold current.

The last regime considered is that of large absolute cooling. If substantial cooling below the heat-sink temperature is desired, simulations with better-optimized structures and packaging have predicted single-stage cooling of 20 to 30°C with cooling power densities of several thousand W/cm^2.[16] This better performance would not only enhance their utility in the above-mentioned applications, but also allow for use in such applications as IR lasers and photodetectors.

6 Conclusions

Monolithically integrated cooling of optoelectronic devices with thin-film solid-state coolers has been demonstrated. Cooling power densities of several hundred W/cm^2 were measured with an integrated InP *pin* diode and a 1-μm-thick InGaAs/InGaAsP HIT cooler. The wire bonds and contact resistance were determined to add significantly to the heat load density and must be minimized in an optimally packaged device. Integration of coolers with diode lasers grown from similar materials is possible, and it should be suitable for significantly controlling the characteristics of the laser sources by manipulating the cooler current. More generally, these monolithically integrated thin-film coolers should have wide applications in any optoelectronic device where it is beneficial to control the operating temperature.

Acknowledgments

This work was supported by the Army Research Office through the DARPA/HERETIC program and by the Office of Naval Research.

References

1. J. Piprek, Y. A. Akulova, D. I. Babic, L. A. Coldren, and J. E. Bowers, "Minimum temperature sensitivity of 1.55 μm vertical-cavity lasers at 30 nm gain offset," *Appl. Phys. Lett.* **72**(15), 1814–1816 (1998).
2. J. Hukriede, D. Kip, and E. Kratzig, "Thermal tuning of a fixed Bragg grating for IR light fabricated in a $LiNbO_3$:Ti channel waveguide," *Appl. Phys. B: Lasers Opt.* **70**, 73–75 (2000).
3. Y. Baek, R. Schiek, and G. I. Stegeman, "All-optical switching in a hybrid Mach-Zehnder interferometer as a result of cascaded second-order nonlinearity," *Opt. Lett.* **20**(21), 2168–2170 (1995).
4. X. Lu, D. An, L. Sun, Q. Zhou, and R. T. Chen, "Polarization-insensitive thermo-optic switch based on multimode polymeric waveguides with an ultralarge optical bandwidth," *Appl. Phys. Lett.* **76**(16), 2155–2157 (2000).
5. J. Piotrowski, C. A. Musca, J. M. Dell, and L. Faraone, "Infrared photodetectors operating at near room temperature," in *Proc. Conf. on Optoelectronic and Microelectronic Materials and Devices*, pp. 124–127, Perth, WA, Australia (1998).
6. L. Rushing, A. Shakouri, P. Abraham, and J. E. Bowers, "Micro thermoelectric coolers for integrated applications," in *Proc. 16th International Conf. on Thermoelectrics*, pp. 646–649, Dresden, Germany (1997).
7. T. A. Corser, "Qualification and reliability of thermoelectric coolers for use in laser modules," in *41st Electronic Components and Technology Conference*, pp. 150–165, Atlanta, GA, USA (1991).
8. A. Shakouri, C. LaBounty, J. Piprek, P. Abraham, and J. E. Bowers, "Thermionic emission cooling in single barrier heterostructures," *Appl. Phys. Lett.* **74**(1), 88–89 (1999).
9. A. Shakouri, E. Y. Lee, D. L. Smith, V. Narayanamurti, and J. E. Bowers, "Thermoelectric effects in submicron heterostructure barriers," *Microscale Thermophys. Eng.* **2**(1), 37–48 (1998).
10. J. M. Swartz and J. R. Gaines, "Wide range thermometry using gallium arsenide sensors," in *Measurement & Control in Science & Industry*, Vol. 4, Harmon H. Plumb, Ed., pp. 1117–1124, Instrument Society of America, Philadelphia, PA (1972).
11. G. K. Reeves and H. B. Harrison, "Obtaining the specific contact resistance from transmission line model measurements," *IEEE Electron Device Lett.* **EDL-3**(5), 111–113 (1982).
12. D. M. Rowe, *CRC Handbook of Thermoelectrics*, CRC Press, New York (1995).
13. N. K. Dutta, T. Cella, R. L. Brown, and D. T. C. Huo, "Monolithically integrated thermoelectric controlled laser diode," *Appl. Phys. Lett.* **47**(3), 222–224 (1985).
14. P. R. Berger, N. K. Dutta, K. D. Choquette, G. Hasnain, and N. Chand, "Monolithically Peltier-cooled vertical-cavity surface-emitting lasers," *Appl. Phys. Lett.* **59**(1), 117–119 (1991).
15. T. S. El-Bawab, C. H. Vaishnav, A. P. Jayasumana, H. Temkin, J. R. Sauer, and H. A. Willebrand, "Robust wavelength division multiplexed local area networks," *Fiber Integr. Opt.* **16**, 237–260 (1997).
16. C. LaBounty, A. Shakouri, G. Robinson, P. Abraham, and J. E. Bowers, "Design of integrated thin film coolers," in *Proc. 18th International Conf. on Thermoelectrics*, Baltimore, MD, USA, 1999.

Christopher LaBounty received his BS in electrical engineering from the University of Vermont in 1997 and his MS in 1999 from the University of California, Santa Barbara. He is currently pursuing his PhD at UCSB, working on heterostructure integrated thermionic coolers and their integration with optoelectronics. His research interests also include optical WDM communication systems and thermal management in optoelectronic components.

Ali Shakouri received his Engineer degree from Ecole Nationale Superieure des Telecommunications de Paris, France, in 1990, and PhD degree from California Institute of Technology in 1995. He is an assistant professor of electrical engineering at the University of California, Santa Cruz. From 1995 to 1998 he was with the Ultrafast Optoelectronics Group at UCSB. His present research is in integrated cooling of optoelectronic devices and photonic switching.

Patrick Abraham received the MS and PhD degrees in material science from the University C. Bernard Lyon 1, France, in 1984 and 1987, respectively. He was a researcher at Centre National de la Recherche Scientifique (CNRS), France, and worked until 1988 at the Laboratoire de Physico-Chimie Minerale, France. He is currently a research engineer at the University of California at Santa Barbara, where his research interests include the study of new active regions for long-wavelength lasers, their growth by MOCVD, and their growth on compliant universal substrates.

John E. Bowers received the MS and PhD degrees from Stanford University, Palo Alto, California. He is the director of the Multidisciplinary Optical Switching Technology Center (MOST), the executive director of the Center for Entrepreneurship and Engineering Management, and a professor in the Department of Electrical and Computer Engineering at the University of California, Santa Barbara (UCSB). He is a member of the Optoelectronics Technology Center and the NSF Science and Technology Center on Quantized Electronic Structures. His research interests are primarily concerned with high-frequency optoelectronic devices and physics. He worked for AT&T Bell Laboratories and Honeywell before joining UCSB in 1987.

SiGe/Si SUPERLATTICE COOLERS

Xiaofeng Fan, Gehong Zeng, Edward Croke[a)], Gerry Robinson, Chris LaBounty, Ali Shakouri[b)], and John E. Bowers

Department of Electrical and Computer Engineering
University of California, Santa Barbara, California 93106, USA
a) HRL Laboratories, LLC, Malibu, California 90265, USA
b) Baskin School of Engineering, University of California, Santa Cruz, California 95064, USA

ABSTRACT

The fabrication and characterization of SiGe/Si superlattice coolers are described. Superlattice structures were used to enhance the device performance by reducing the thermal conductivity between the hot and the cold junctions, and by providing selective removal of hot carriers through thermionic emission. Cooling of 2.2 K and 2.5 K were measured on n-type and p-type 75 × 75 μm^2 devices, corresponding to cooling power densities of hundreds of watts per square centimeter. Cooling up to 7.2 K was obtained at 150 °C for p-type 50 × 50 μm^2 devices. The results show that n-type and p-type coolers can work together in similar optimal conditions. This paves the road to fabricate n-type and p-type superlattice coolers in an array format electrically in series and thermally in parallel, similar to conventional thermoelectric devices, and thus achieve large cooling capacities with relatively small currents.

INTRODUCTION

The demand for high-speed, high-density very-large-scale integrated (VLSI) circuits is accompanied by higher power densities. Many devices are already operating at or near the edge of their reliability. Heat generation and thermal management are becoming one of the barriers to further increase clock speeds and decrease feature sizes. Thermoelectric (TE) refrigeration is a solid-state active cooling method with high reliability. TE coolers are silent, environmentally "green" and capable for spot cooling. Bi_2Te_3 based TE coolers are commonly used in electronics and optoelectronics for cooling and temperature stabilization, but their manufacturing is a bulk technology and is incompatible with integrated circuit (IC) fabrication process. Solid-state coolers monolithically integrated with VLSI devices are an attractive way to achieve compact and efficient cooling. However, the TE figure of merit (Z) is quite low for most of the semiconductors used in microelectronics and optoelectronics. This makes it difficult to get high performance integrated coolers. Recently heterostructure thermionic coolers and superlattice coolers are proposed, and theoretical calculations show that large improvement in Z can be achieved and efficient cooling becomes possible with coolers made of conventional semiconductor materials [1-6]. InP based heterostructure integrated thermionic coolers have been demonstrated by Shakouri et al, one degree cooling was obtained over 1 μm InGaAsP barrier [7-8]. In this paper, we report our experimental results for SiGe/Si superlattice coolers.

SiGe is a good thermoelectric material for high temperature refrigeration and power generation applications [9]. It has been used for thermo-nuclear power generation in satellites for deep space missions. In this paper we describe the fabrication and characterization of single-element SiGe/Si superlattice coolers that use thermionic and thermoelectric effects. Both n-type and p-type devices have been demonstrated. This paves the road to fabricate n-type and p-type superlattice coolers in an array format electrically in series and thermally in parallel, similar to conventional thermoelectric devices, and thus achieve large cooling capacities with relatively small currents. Superlattice structures can enhance the cooler performance by reducing the thermal conductivity between the hot and the cold junctions [10] and by selective emission of hot carriers above the barrier layers in the thermionic emission process [1-2, 7]. Si and SiGe-based devices can be monolithically integrated with these coolers to achieve better device performance.

MATERIAL AND DEVICE FABRICATION

SiGe/Si superlattice structures were grown in a Perkin-Elmer Si molecular beam epitaxy (MBE) growth chamber on 125 mm diameter, (001)-oriented Si substrates doped to <0.020 Ω-cm with Sb for the n-type devices and to <0.006 Ω-cm with B for the p-type devices.

The cooler's main part is a 3 μm thick $Si_{0.7}Ge_{0.3}$/Si superlattice grown symmetrically strained on a buffer layer designed so that the in-plane lattice constant approximates that of relaxed $Si_{0.9}Ge_{0.1}$. The doping levels are 2×10^{19} cm^{-3} and 5×10^{19} cm^{-3} for n-type and p-type SiGe/Si superlattices, respectively.

For the relaxed buffer layer, we grew a 10-layer structure, alternating between 150 nm $Si_{0.9}Ge_{0.1}$ and 50 nm $Si_{0.845}Ge_{0.150}C_{0.005}$, roughly following the method suggested by Osten et al. [11]. For the n-type sample, the layers were grown at 390 °C and annealing was performed at 750 °C for 10 minutes after the growth of each $Si_{0.9}Ge_{0.1}$ layer. In the p-type case, the growth temperature was simply alternated between 700 °C for the $Si_{0.9}Ge_{0.1}$ layer and 500 °C for the $Si_{0.845}Ge_{0.150}C_{0.005}$ layer. After the relaxed buffer sequence, another 150 nm thick $Si_{0.9}Ge_{0.1}$ layer was grown at 390 °C for the n-type sample and a 1 μm thick $Si_{0.9}Ge_{0.1}$ layer was grown at 700 °C for the p-type case. Growth of a 200 period, 5 nm $Si_{0.7}Ge_{0.3}$/ 10 nm Si superlattice then followed at 390 °C (n-type case) and 500 °C (p-type case). Finally, the samples were capped with a heavily doped $Si_{0.9}Ge_{0.1}$ layer sequence to provide for a low-resistance ohmic contact.

The processing of SiGe/Si superlattice coolers is compatible with that of VLSI technology. Mesas 3.6 μm high were fabricated using reactive ion etching down to the SiGe buffer layer to form the devices. Metallization was made on the mesa and $Si_{0.9}Ge_{0.1}$ buffer layer for top and bottom contact respectively. A scanning electron microscope (SEM) image of the processed devices is shown in Fig. 1.

Electrical contact resistance is an important factor that limits the optimum device performance. Low contact resistance is essential for thin film coolers [12]. A 100 nm titanium metal layer was first deposited. This was intended to form a titanium silicide on the silicon surface and to act as a metal barrier to separate Si and Al. Subsequently 1 μm thick aluminum layer was deposited. To facilitate wire bonding, an additional metal layer of titanium and gold was used. Annealing was accomplished at temperatures between 400 °C and 600 °C with rapid thermal annealer. TLM (transmission line mode) measurements were carried out to measure the contact resistance. The measured specific contact resistivity is in the mid 10^{-7} Ω-cm^2 range for both n-type and p-type devices as shown in Fig. 2.

TEST RESULTS AND DISCUSSIONS

As shown in Fig. 3, the SiGe/Si superlattice coolers were tested on a temperature controlled copper plate that worked as the heat sink. The heat sink is set at a constant temperature during the device testing. To cool the top of the devices, current was sent from the top metal contact to bottom metal contact for p-type device and the reverse direction for n-type device. Micro thermocouples were used to measure the cooling temperatures. Fig. 4 displays the measured temperature on top of the 75 × 75 μm^2 n-type and p-type devices as a function of current with the heat sink at 25 °C. The cooling temperature is relative to the value at zero current. Despite the large thermal resistance of the Si substrate and package on the hot side of the cooler and Joule heating in the wires connected to the cold junction, a net cooling of 2.2 K and 2.5 K was observed on top of the n-type and p-type devices respectively. This cooling over the small barrier thickness corresponds to cooling capacities on the order of hundreds of watts per square centimeter.

Devices of different sizes, from 50 × 50 μm^2 to 150 × 150 μm^2, were tested. The results on p-type SiGe/Si coolers are shown in Fig. 5. 2.7 K cooling is obtained for the 50 × 50 μm^2 p-type device. The test results shows the maximal cooling temperature increases as the device size decreases. This cannot be explained with conventional ideal thermoelectric or thermionic cooler models. This is due to the three-dimensional (3D) nature of current spreading in the substrate and

the Joule heating from the bonding wires. For the same wire resistance, smaller devices require a smaller optimum current which is favorable for better cooling performance. A 3D finite-difference heat equation solver [8] is being used to model the device performance.

The SiGe/Si superlattice coolers have a better performance at higher temperatures. The measured cooling for 50 × 50 μm^2 p-type SiGe/Si devices at 150 °C (heat sink temperature) is shown in Fig. 6. The net cooling increases from 2.7 K at 25 °C to 7.2 K at 150 °C. The reason for improved performance with the increase in temperature is two fold. First, in the temperature range of our measurements, the figure of merit ZT of SiGe alloy increases with temperature due to smaller thermal conductivity and larger Seebeck coefficient at higher temperatures [13], and second, the thermionic emission cooling power increases due to the larger thermal spread of carriers near the Fermi energy.

Since the devices described above are single element superlattice coolers, heat conduction to the cooling side from the bonding wires or probes are unavoidable. This reduces the maximum cooling. To solve this problem, n-type and p-type SiGe/Si superlattice coolers can be made in an array format electrically in series and thermally in parallel, similar to conventional thermoelectric coolers. In this way, both electrical terminals can be made at the heat sink side, and large cooling capacities can be achieved with relatively small currents. For the cooler array, it is preferable to have n-type and p-type devices work at the same optimal current. This can be done by selecting suitable device sizes for the cooler couple. For example, optimal operation currents for the n-type 75 × 75 μm^2 devices (in Fig. 4) and the p-type 50 × 50 μm^2 devices (in Fig. 5) are about the same, cooling of 2.2 K and 2.7 K can be obtained at the same current ~230 mA.

With optimized superlattice material and device design and packaging, cooling up to tens of degrees is possible. More important, the processing of SiGe/Si superlattice coolers is compatible with that of VLSI technology, thus it is possible to integrate these coolers monolithically with Si and SiGe devices to achieve compact and efficient cooling.

CONCLUSIONS

SiGe/Si superlattice coolers were demonstrated and cooling of 2.2 K and 2.5 K were measured for n- and p-type 75 × 75 μm^2 devices. Cooling up to 7.2 K was obtained at 150 °C for p-type 50 × 50 μm^2 devices. The results show that the packaged devices of both n- and p-type coolers can work together with similar bias current conditions. This paves the road to fabricate n- and p-type superlattice coolers in an array format electrically in series and thermally in parallel, similar to conventional thermoelectric devices. Thus large cooling capacities with relatively small currents can be achieved and the problems of series resistance and heat load of contacting wires can be avoided.

ACKNOWLEDGMENTS

This work was supported by the DARPA HERETIC program and the Army Research Office. X. F. would like to acknowledge many stimulating discussions with Professor Venky Narayanamurti.

REFERENCES

[1] A. Shakouri, J.E. Bowers, *Appl. Phys. Lett.*, **71** (1997), 1234.

[2] A. Shakouri, C. Labounty, P. Abraham, J. Piprek, and J.E. Bowers, in *Material Research Society Symposium Proceedings*, **545** (1999), 449.

[3] L.K. Hicks, M.S. Dresselhaus, *Phys. Rev. B*, **47** (1993), 12727.

[4] L.D. Hicks, T.C. Harman, M.S. Dresselhaus, *Appl. Phys. Lett.*, **63** (1993), 3230.

[5] T. Koga, X. Sun, S.B. Cronin, M.S. Dresselhaus, *Appl. Phys. Lett.*, **73** (1998), 2950.

[6] T. Koga, X. Sun, S.B. Cronin, M.S. Dresselhaus, *Appl. Phys. Lett.*, **75** (1999), 2438.

[7] A. Shakouri, P. Abraham, C. LaBounty, J.E. Bowers, in *Proceedings of the 17th International Conference on Thermoelectrics*, 1998, P. 218.

[8] A. Shakouri, C. LaBounty, J. Piprek, P. Abraham, J.E. Bowers, *Appl. Phys. Lett.*, **74** (1999), 88.

[9] C. B. Vining, *J. Appl. Phys.*, **69** (1991), 331.

[10] S.-M Lee, D. G. Cahill, R. Venkatasubramanian, *Appl. Phys. Lett.*, **70** (1997), 2957.

[11] H. J. Osten, E. Bugiel, *Appl. Phys. Lett.*, **70** (1997), 2813.

[12] C. LaBounty, A. Shakouri, G. Robinson, P. Abraham, J.E. Bowers, in *Proceedings of the 18th International Conference on Thermoelectrics*, Baltimore, MD, USA, 1999.

[13] J.P. Dismukes, L. Ekstrom, E.F. Steigmeier, I. Kudman, D.S. Beers, *J. Appl. Phys.*, **35** (1964), 2899.

FIGURE CAPTIONS

Figure 1. SEM image of the processed SiGe/Si superlattice cooler devices.

Figure 2. Ohmic contact resistance for n-type and p-type SiGe based on TLM measurements. The doping is 2×10^{20} cm^{-3} for p-type SiGe and 1×10^{20} cm^{-3} for n-type SiGe.

Figure 3. Schematic diagram of SiGe/Si superlattice cooler for testing (not to scale).

Figure 4. Measured cooling for n-type and p-type 75 × 75 μm^2 SiGe/Si superlattice coolers at 25 °C heat sink temperature.

Figure 5. Cooling measured on top of p-type SiGe/Si superlattice coolers of different sizes at 25 °C heat sink temperature.

Figure 6. Measured cooling for a 50 × 50 μm^2 p-type SiGe/Si superlattice cooler at 150 °C heat sink temperature.

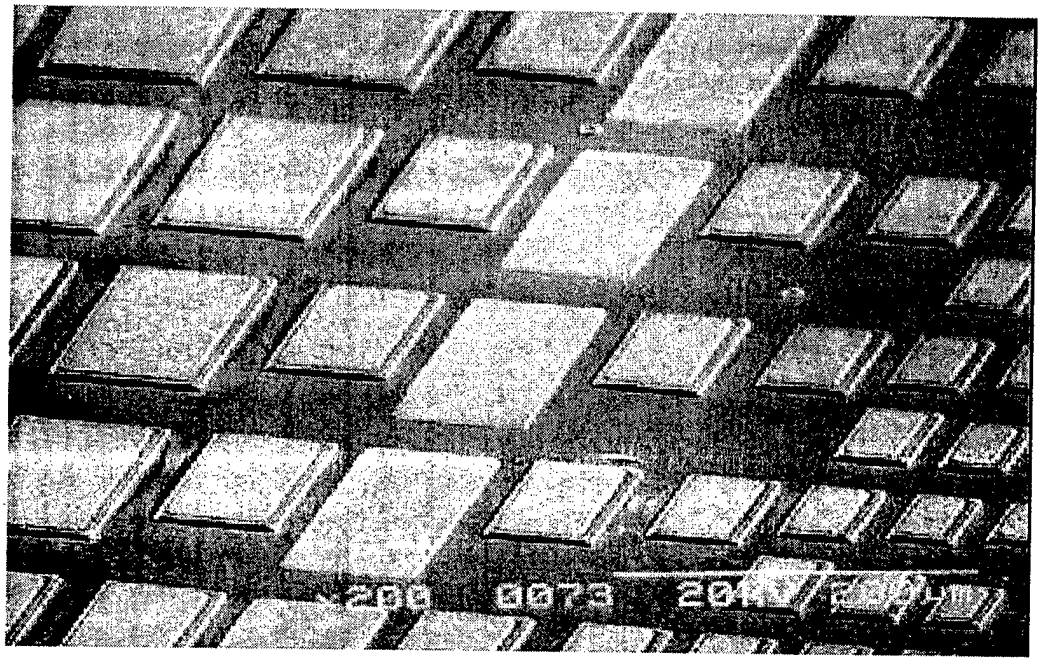


Fig. 1

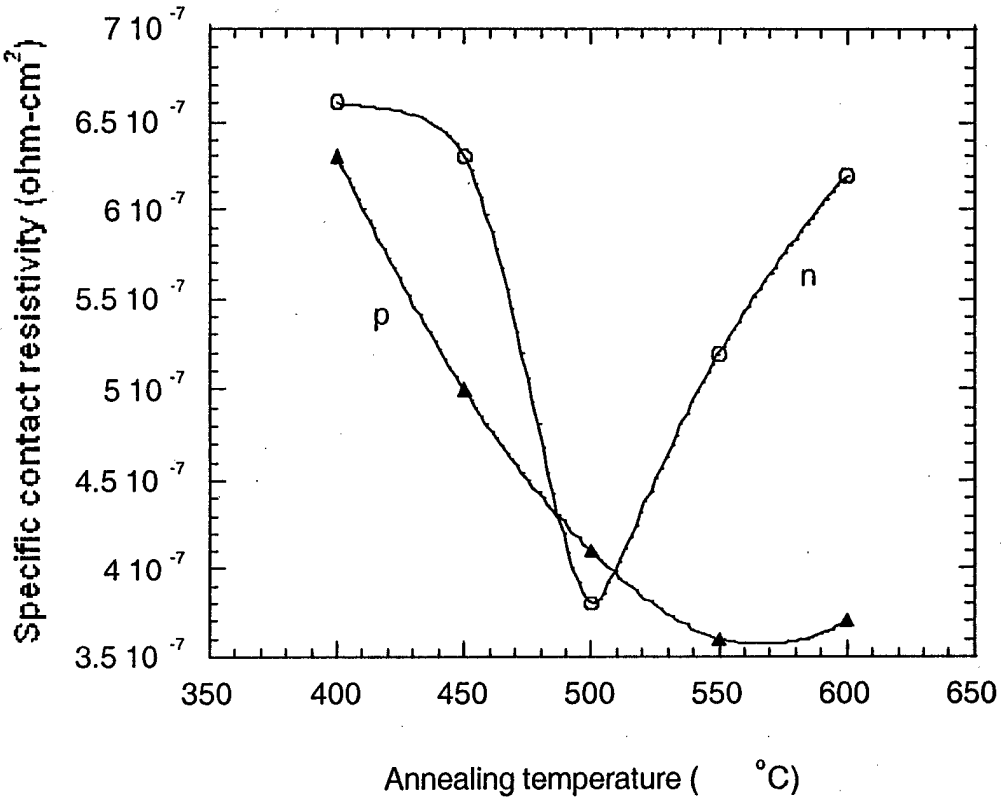

Specific contact resistivity (ohm-cm^2)
7 10^{-7}
6.5 10^{-7}
6 10^{-7}
5.5 10^{-7}
5 10^{-7}
4.5 10^{-7}
4 10^{-7}
3.5 10^{-7}
350
400
450
500
550
600
650
Annealing temperature (°C)
p
n

Fig. 2

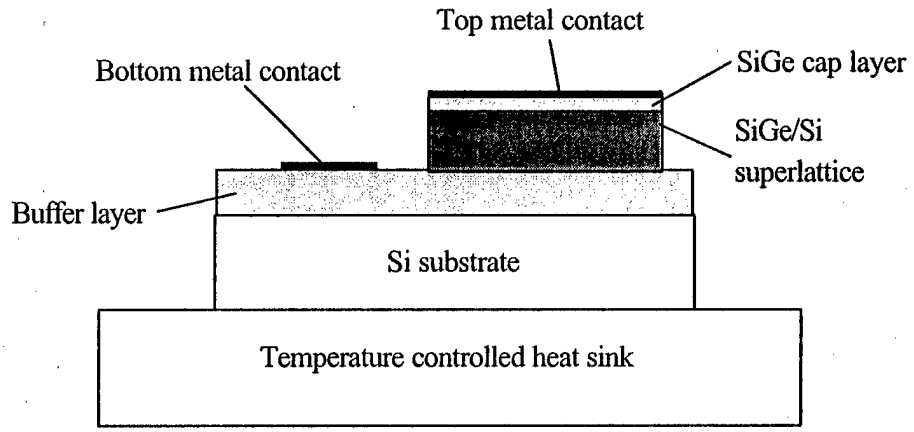

Top metal contact
Bottom metal contact
SiGe cap layer
SiGe/Si
superlattice
Buffer layer
Si substrate
Temperature controlled heat sink

Fig. 3

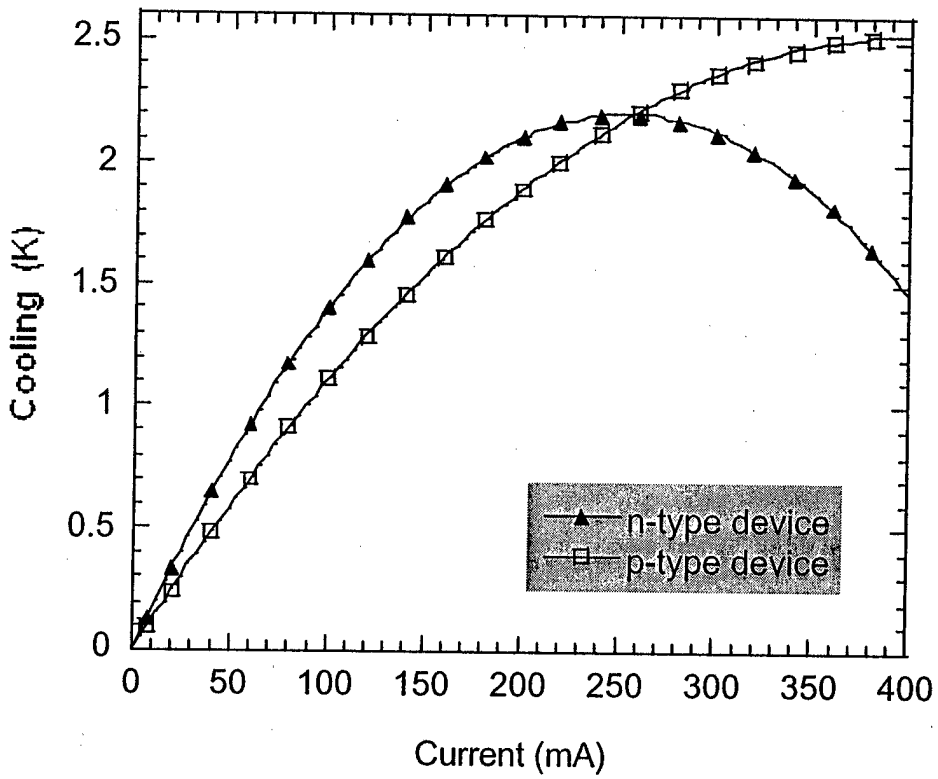

2.5
2
1.5
1
0.5
0
Cooling (K)
0
50
100
150
200
250
300
350
400
Current (mA)
n-type device
p-type device

Fig. 4

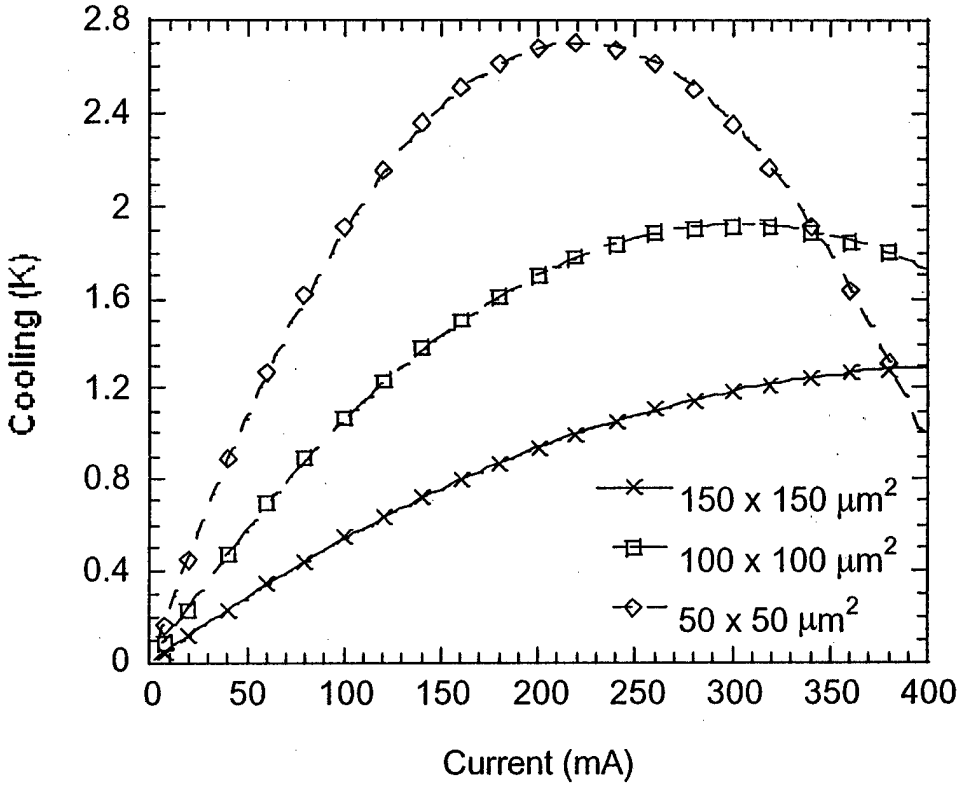

2.8
2.4
2
1.6
1.2
0.8
0.4
0
Cooling (K)
0
50
100
150
200
250
300
350
400
Current (mA)
150 x 150 μm²
100 x 100 μm²
50 x 50 μm²

Fig. 5

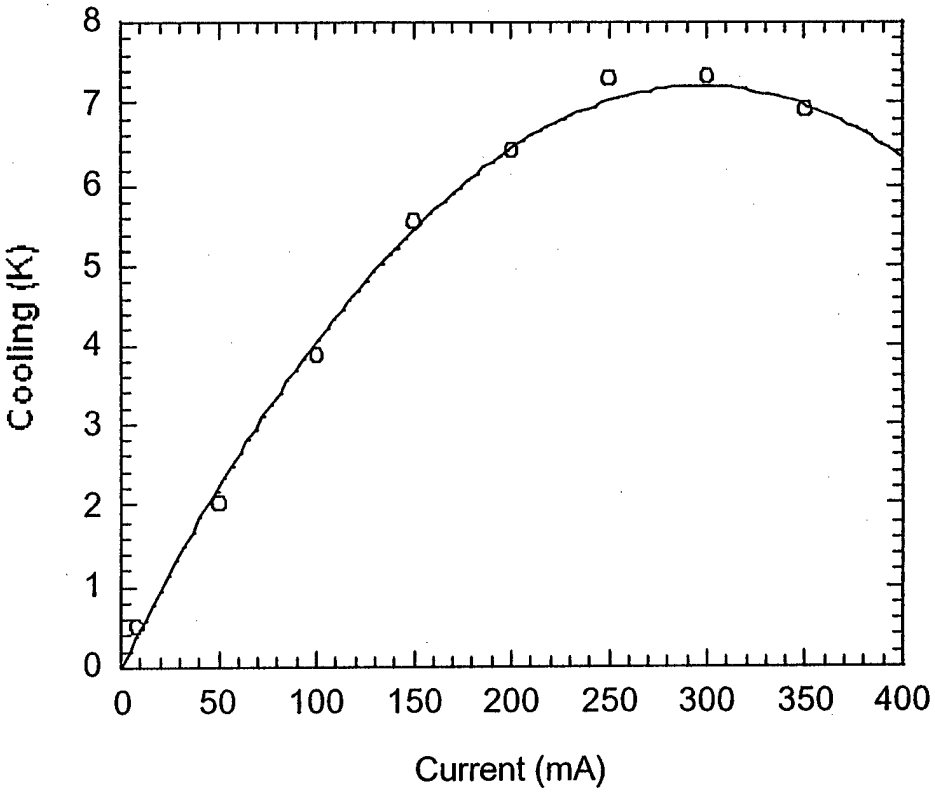

Cooling (K)
0
1
2
3
4
5
6
7
8
0
50
100
150
200
250
300
350
400
Current (mA)

Fig. 6

Thermal Conductivity of Indium Phosphide Based Superlattices

Scott Huxtable (1), Ali Shakouri (2), Patrick Abraham (3), Yi-Jen Chiu (3), Xiafeng Fan (3), John E. Bowers (3), Arun Majumdar* (1)

(1) Department of Mechanical Engineering
University of California - Berkeley
Berkeley, California 94720, USA

(2) Jack Baskin School of Engineering
University of California - Santa Cruz
Santa Cruz, CA 95064-1077, USA

(3) Department of Electrical and Computer Engineering
University of California - Santa Barbara
Santa Barbara, CA 93106-9560, USA

Abstract

Semiconductor superlattice structures have shown promise as thermoelectric materials for their high power factor and low thermal conductivity [1,2]. Heat conduction by phonons in semiconductor superlattices is controlled by two mechanisms: interface scattering by defects and/or roughness and phonon filtering by Bragg reflection. The thermal conductivity of three different InP/InGaAs superlattices was measured from 77 - 300 K using the 3ω method. Small "dips" in the measured thermal conductivity were observed near 90 K, which may be indicative of phonon filtering.

Introduction

The thermal conductivity of a thermoelectric material is a key parameter in both the thermoelectric (TE) and thermionic (TI) figure of merit. In both cases, the device performance increases with a decrease in thermal conductivity. Heat transport in semiconductors is dominated by phonons for carrier densities below 10^{20} cm^{-3}. Recent work has shown that the thermal conductivity of semiconductor superlattices is often lower than the value as calculated from their constituent materials using Fourier heat conduction theory [3-8]. Additionally, in some instances, thermal conductivities have been reported that are below the value of a corresponding alloy [5-8]. This reduction of thermal conductivity has mainly been attributed to two mechanisms: (i) phonon scattering at interfaces due to roughness, defects, and/or dislocations, (ii) phonon "filtering" [9]. Previous measurements could not separate out these two effects because the superlattices were not lattice matched. It is well known that interface scattering will reduce the phonon mean free path and, hence, decrease the heat transport. However, an increase in scattering sites also increases the electrical resistance causing an undesirable drop in the power factor $S^2\sigma$. Therefore, it seems that the reduction of thermal conductivity through the use of phonon bandgaps and phonon filtering should be explored.

In this work we measured the thermal conductivity of three different lattice matched InGaAs/InP superlattices. The reasons for studying this particular superlattice system are twofold. One is that InGaAs/InP superlattices are currently being utilized as thermionic coolers [10]. Additionally, these superlattices are lattice matched with high quality interfaces, such that the effects of the interface scattering will be greatly diminished. For this reason it may be possible to separate out the effects of phonon filtering and interface scattering.

Analogous to optical filtering by distributed Bragg reflection, phonon filtering takes place when phonons of wavelength λ_o satisfy the Bragg condition $\lambda_o=2d_o$, where d_o is the superlattice period. In a very nice set of experiments, Narayanamurti et al. [9] showed that for a GaAs/AlGaAs superlattice, longitudinal phonon transmission near one meV can be reduced by roughly 60 percent over a bandgap of about 0.1 meV. However, that experiment was done with monochromatic phonons, while heat is conducted over a wide range of phonon frequencies. Therefore, it is more difficult to observe the phonon filtering phenomenon directly from thermal conductivity data. However, since the dominant phonon frequency is dependent on temperature it is possible to study spectral filtering through the Planck blackbody distribution by measuring the thermal conductivity over a range of temperatures.

Experimental Details

The thermal conductivity of the InGaAs/InP superlattices was measured using the 3ω technique [7]. The 3ω method is an ac technique that utilizes a narrow metal line that is patterned on the surface of the sample as both a heater and a

* Corresponding author: majumdar@me.berkeley.edu

thermometer. A current with angular frequency ω is passed through the metal line, which heats the surface of the sample through Joule heating at a frequency 2ω. The resistance of a pure metal increases with increasing temperature, therefore causing the electrical resistance of the heater to also oscillate at 2ω. This 2ω component times the original driving current at ω produces a small voltage oscillation at 3ω which can be used to determine the thermal response of the sample.

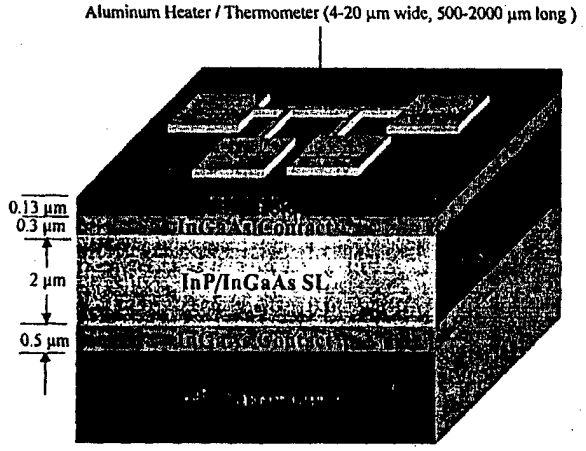


Figure 1: Superlattice with aluminum heater patterned on top. The two InGaAs layers are used as electrical contacts for other measurements, while the PECVD oxide is deposited on top for electrical isolation between the aluminum heater and the superlattice.

Figure 1 shows a cross-sectional view of the sample. These samples were also used in a variety of other experiments and the InGaAs layers above and below the superlattice serve as contacts. The lattice-matched InP/InGaAs superlattice was deposited by MetalOrganic Chemical Vapor Deposition (MOCVD). For a dielectric sample the metal heater/thermometer line may be placed directly on the surface. However, since the superlattices are electrically conducting a thin insulating layer must be deposited on the sample first. We used a 130 nm thick SiO_2 layer that is deposited using Plasma Enhanced Chemical Vapor Deposition (PECVD) at 300°C. An aluminum heater with a thickness of ~200 nm was then evaporated and patterned on top of the oxide layer using the lift off technique. The heaters used in this work were typically 5 μm wide and 500 μm long.

The temperature rise at the surface was measured over a wide range of frequencies. This measurement gave the series resistance of the substrate and the films on the surface. The thermal conductivity of the substrate was readily obtained from the slope of the measured temperature rise (ΔT) vs. the natural log of frequency using [11]

$$K_{sub} = \frac{V^3 \ln(f_2/f_1)}{4\pi l R^2 (V_{3\omega,1} - V_{3\omega,2})} \frac{dR}{dT} \qquad (1)$$

where K_{sub} is the thermal conductivity of the substrate, V is the voltage across the metal line at ω, R is the resistance of the metal line, l is the length of the metal line, f_2 and f_1 are the two frequencies used to find the slope, $V_{3\omega,2}$ and $V_{3\omega,1}$ are the in phase 3ω voltages at the two frequencies, and dR/dT is the measured resistance change with temperature of the metal line.

Once the thermal conductivity is known, the ΔT of the substrate can be calculated [12]. The thermal conductivity of the InP substrate can be found in Figure 4. Since each film is thin and has little capacitance, it simply adds a frequency independent ΔT to the temperature rise of the substrate.

$$\Delta T_{Total} = \Delta T_{Sub} + \Delta T_{InGaAs} + \Delta T_{SL} + \Delta T_{Oxide} \qquad (2)$$

The thermal conductivities for the InGaAs and oxide layers were measured separately on other samples using [12]

$$K_{film} = \frac{Pt}{\Delta T_{film} wl} \qquad (3)$$

where K_{film} is the film thermal conductivity, P is the heater power, t is the film thickness, ΔT_{film} is the temperature rise of the film, and w and l are the width and length of the heater/thermometer line, respectively. These measurements were performed on separate samples that accompanied the superlattices through the PECVD oxidation, photolithography, and lift off processes. These samples were 130 nm of PECVD oxide on a bare silicon wafer and 2 μm of InGaAs on an InP substrate with the 130 nm SiO_2 layer on top. The thermal conductivity of the InGaAs and oxide are shown in Figures 3 and 4, respectively. Once the contributions of the InGaAs and oxide layers were determined, they were subtracted out from the raw data, which left only the ΔT of the superlattice. The thermal conductivity of the superlattice was then determined from eqn (3) in the same manner as for the InGaAs and oxide.

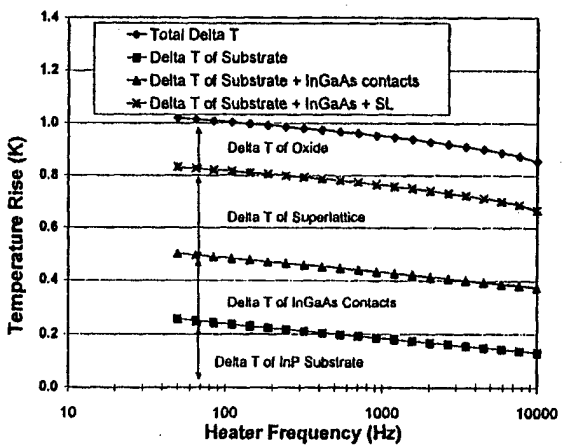


Figure 2: Temperature oscillations as a function of frequency for a superlattice. Note the frequency independent temperature rise associated with each thin film.

The experiments were conducted in a liquid nitrogen cryostat over a temperature range of 77 - 320 K and the data was taken with a lock in amplifier.

Results and Discussion

Three InGaAs/InP superlattices were measured in this study. All three samples had 80 superlattice periods with each period 25 nm thick to give a total thickness of 2μm. The ratio of the thickness of the InGaAs layer to the InP layer in each period

was different for the three samples. The first superlattice contained 18 nm of InGaAs and 7 nm of InP in each period, while for the 2nd and 3rd samples this ratio was 20:5 and 22:3, respectively. This ratio was changed in order to shift the location of the phonon bandgaps.

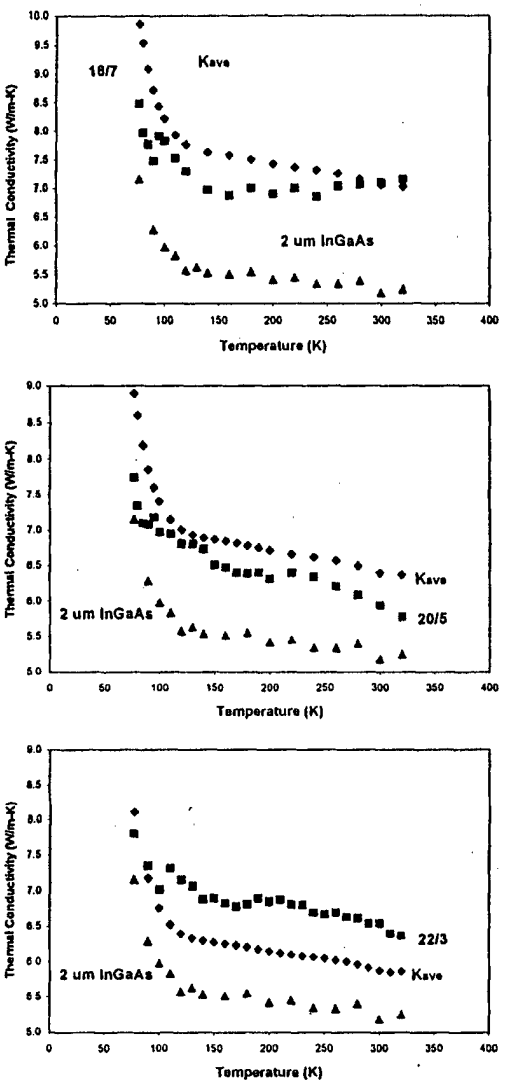


Figure 3: Thermal conductivities for the three superlattices. The labels refer to the ratio of InGaAs to InP in each period (i.e. 18/7 is for the superlattice with a period of 18 nm of InGaAs and 7 nm of InP). All superlattices had 80 periods of 25 nm each giving a total thickness of 2 μm.

In this study, only relatively narrow lines (~5 μm) were used. Since the half width of the line is approximately the same as the film thickness, thermal conductivities for the superlattices and the InGaAs film reported here are a combination of the in plane and cross plane values. In the future, wide lines will be used in conjunction with the narrow ones along with a finite difference model of the heat conduction in order to extract the two values [13].

Figure 3 shows the measured thermal conductivities for the three superlattices studied. The 2 μm InGaAs layer is included for reference. K_{ave} is simply an "average" thermal conductivity that was calculated also for reference. It is calculated from eqn (4) which does not take into account any interface scattering or phonon interactions.

$$K_{ave}=(d_1+d_2)K_1K_2/(K_2d_1+K_1d_2) \quad (4)$$

where d is the thickness of each layer, K is the thermal conductivity of the layer, and the subscripts 1 and 2 refer to InGaAs and InP, respectively.

The uncertainty in the absolute value of the thermal conductivities is rather high at about 25 percent. This uncertainty was dominated by the uncertainty in the widths of the aluminum lines. Despite the fact that that the widths were measured with an Atomic Force Microscope (AFM), there was some variation in the line width along the length of the heater/thermometer which left us with an uncertainty of approximately ± 0.5 μm. This accounted for over half of the total overall uncertainty. The large uncertainty is likely the reason that K_{ave} was above the thermal conductivity for two of the superlattices, but below the third.

However, it is important to note that this uncertainty was constant over the temperature range studied. Therefore, while the absolute value of the thermal conductivity has a high uncertainty, we expect that the general trends of the data are accurate. With that in mind, it is apparent that for all three superlattices measured there is a slight, but noticeable, "dip" in the thermal conductivity curve between 90 and 100 K. The exact reason for this feature is unclear at the present time, although we suggest that it could be indicative of phonon filtering at this temperature.

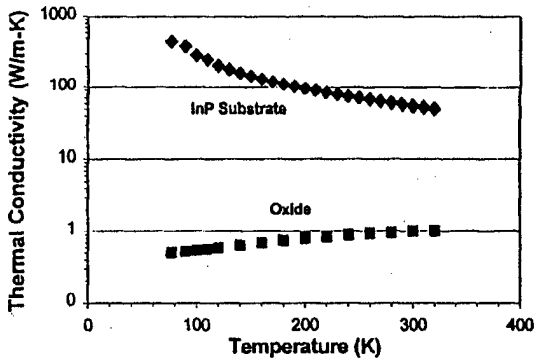


Figure 4: Measured thermal conductivities for the InP substrates and the 130 nm PECVD SiO_2. These values were used to determine the temperature rise in each film for eqn (3).

Future Work

Clearly there is more work to be done. The finite difference model will be completed to extract K_{normal} and $K_{parallel}$. Additionally, variations in the number of periods will be used in order to gain insight as to the role interface scattering in heat conduction.

Acknowledgments

This work was supported by the DARPA Heretic program with Elias Towe as the program director and the Office of Naval Research. S. H. would like to thank Gerry Robinson

and Minyao Mao for helpful discussions regarding processing, David Cahill for advice with the 3ω method, and Li Shi for measurements on the AFM.

References

1. Venkatasubramanian, R., Colpitts, T., Mat. Res. Soc. Symp. Proc., Vol. 478, p. 73, (1997).
2. Hicks, L. D., Harmon, T. C., Sun, X., and Dresselhaus, M. S., Phys. Rev. B, Vol. 53, R10 493 (1996).
3. Yao, T., Appl. Phys. Lett. Vol. 51, p. 1798 (1987).
4. Yu, X. Y., Chen, G., Verma, A., and Smith, J., S., Appl. Phys. Lett., Vol. 67, p. 3553 (1995).
5. Capinski, W. S., and Maris, H. J., Physica B, Vol. 219, 220, p. 699 (1996).
6. Venkatasubramanian, R., Naval Research News, XLVIII, p. 31-43 (1996).
7. Lee, S.-M. and Cahill, D. G., Appl. Phys. Lett., Vol. 70, p. 2957 (1997).
8. Chen, G., Tien, C. L., Wu, X., Smith, J. S., J. Heat Transfer, Vol. 116, p. 325, (1994).
9. Narayanamurti, V., Störmer, H. L., Chin, M. A., Gossard, A. C., Wiegmann, W., Phys. Rev. Lett., Vol. 43, No. 27, p. 2012 (1979).
10. Shakouri, A., Bowers, J. E., Appl. Phys. Lett., Vol. 71, no. 9, p. 1234 (1997).
11. Cahill, D., G., Rev. Sci. Instrum. Vol. 61, p. 802 (1990).
12. Lee, S. M., and Cahill, D., G., J. Appl. Phys., Vol. 81, No. 6, p. 2590 (1997).
13. Chen, G., Zhou, S. Q., Yao, D.-Y., Kim, C. J., Zheng, X. Y., Liu, Z. L., Wang, K. L., Proc. Of the 17th International Conference on Thermoelectrics, p. 202 (1998).